

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 10, 2004, 09:18:38 ; Search time 6129.06 Seconds
(without alignments)
18036.975 Million cell updates/sec

Title: US-10-084-817-30

Perfect score: 3702

Sequence: 1 gcctaggggaatgacagcat.....aaggggcggttaccgagaac 3702

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: em_estda:*
2: em_esthda:*
3: em_estnu:*
4: em_estnu:*
5: em_estnu:*
6: em_estnu:*
7: em_estnu:*
8: em_estnu:*
9: em_estnu:*
10: em_estnu:*
11: em_estnu:*
12: em_estnu:*
13: em_estnu:*
14: em_estnu:*
15: em_estnu:*
16: em_estnu:*
17: em_estnu:*
18: em_estnu:*
19: em_estnu:*
20: em_estnu:*
21: em_estnu:*
22: em_estnu:*
23: em_estnu:*
24: em_estnu:*
25: em_estnu:*
26: em_estnu:*
27: em_estnu:*
28: em_estnu:*
29: em_estnu:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	822.4	22.2	923	13	BQ718603 AGENCOURT
2	809.8	21.9	1201	9	AL577531 AL577531
3	790.6	21.4	826	14	CK229986 ILUMINIGEN
4	769	20.8	941	13	BQ925009 AGENCOURT

Result No.	Score	Match	Length	ID	Description
5	768.2	20.8	846	13	BUI51076 AGENCOURT
6	768	20.7	881	14	CD359921 AGENCOURT
7	753.8	20.4	827	14	CD643635 AGENCOURT
8	747.4	20.2	749	9	AL700988 DKEP2686D
9	737.8	19.9	813	13	AUI32922 AGENCOURT
10	732	19.8	823	14	CD656598 AGENCOURT
11	725.4	19.6	745	12	BQ010124 AGENCOURT
12	719.6	19.4	753	9	A1926727 wos0901.x
13	708.8	19.1	753	9	A1926727 wos0901.x
14	708.8	19.1	753	9	A1926727 wos0901.x
15	701.2	18.9	735	13	BX642011 AGENCOURT
16	698.4	18.9	797	14	CB998236 AGENCOURT
17	692.6	18.7	699	12	BQ014967 UI-H-ED1-
18	691.8	18.7	704	13	BX472022 DKEP2686M
19	688.8	18.6	700	13	BQ209023 RST8536
20	688.4	18.6	834	13	BUI567576 AGENCOURT
21	681.4	18.4	741	14	CD643241 AGENCOURT
22	680.8	18.4	887	10	BE749163 AGENCOURT
23	678	18.3	760	14	CB229386 AGENCOURT
24	671.2	18.1	727	13	BX504392 DKEP2686O
25	668.2	18.0	704	9	AUI39912 AGENCOURT
26	666.2	18.0	727	13	AUI39912 AGENCOURT
27	665.2	18.0	727	13	AUI39912 AGENCOURT
28	663	17.9	1201	9	AL554932 AGENCOURT
29	662.2	17.9	840	14	CB310135 AGENCOURT
30	662	17.9	730	13	BX504360 DKEP2686M
31	659.4	17.8	677	13	BX611260 DKEP2686N
32	654.6	17.7	675	13	BX481847 DKEP2686N
33	652.4	17.6	660	13	BX471556 DKEP2686C
34	650.4	17.6	662	12	BQ014498 UI-H-ED1-
35	650.4	17.6	662	12	CA440470 UI-H-ED1-
36	650.2	17.6	655	13	BX476085 DKEP2686E
37	649.4	17.5	651	13	BX479376 DKEP2686E
38	647.8	17.5	664	13	BX505684 DKEP2686H
39	647.2	17.5	662	14	CA432157 UI-H-FG1-
40	645.6	17.4	662	13	BUI565233 UI-H-FG1-
41	641.8	17.3	675	9	AUI39864 AGENCOURT
42	641	17.3	663	12	BM723358 UI-E-ED1-
43	639.4	17.2	851	14	CD251635 AGENCOURT
44	637.6	17.2	759	14	CB962709 AGENCOURT
45	636	17.2	646	14	CA771730 i082811.x

ALIGNMENTS

RESULT 1
BQ718603
LOCUS
DEFINITION
AGENCOURT_8295719 lupo1_sympathetic_trunk Homo sapiens cDNA clone
IMAGE:6188951.5, mRNA sequence.
BQ718603
AGENCOURT_8295719 GI:21857500
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
NIH-MGC http://mgc.nci.nih.gov/
1 (bases 1 to 923)
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Consortium (LNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLM1358 row: f column: 24
High quality sequence stop: 583.

FEATURES

Source

Location/Qualifiers

1. 923
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6188951"
 /sex="male"
 /tissue_type="sympathetic trunk"
 /dev_stage="adult, 16 yr"
 /lab_host="DH10B"
 /clone_lib="Lupski_sympathetic_trunk"
 /note="Vector: pCMV-SPORT6 (Life Technologies); Site 1:
 NotI; Site 2: SalI; cDNA made by oligo-dT priming.
 Directionally cloned using the following adaptors:
 5'-TCGACCCAGGCGTCCG-3' and
 5'-GACTAGTCTTACATATCGCAGCGCGCGCT(15)-3'. Size selected >
 1 kb for average insert length 1.9 kb. This is a primary
 library, non-amplified. Library constructed by Life
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
 College of Medicine); available through Life
 Technologies."

ORIGIN

Query Match 22.2%; Score 822.4; DB 13; Length 923;
 Best Local Similarity 95.9%; Pred. No. 1.9e-132;
 Matches 885; Conservative 0; Mismatches 34; Indels 4; Gaps 4;

QY 492 AGGAGATGCGACTCGAAGCAGAAAGATCGCCTACCGAAGAAATGACGATGATGAAGAGG 551
 1 AGGAGATGCGACTCGAAGCAGAAAGATCGCCTACCGAAGAAATGACGATGATGAAGAGG 60
 QY 552 AGGAGATGCGACTCGAAGCAGAAAGATCGCCTACCGAAGAAATGACGATGATGAAGAGG 611
 1 AGGAGATGCGACTCGAAGCAGAAAGATCGCCTACCGAAGAAATGACGATGATGAAGAGG 60
 Db 61 AGGAGATGCGACTCGAAGCAGAAAGATCGCCTACCGAAGAAATGACGATGATGAAGAGG 120
 QY 612 AAGATCCTTGGGACAGATGACCGACCAAGTGGAGTGAATGCCCAAGACAGTGTGCTTG 671
 121 AAGATCCTTGGGACAGATGACCGACCAAGTGGAGTGAATGCCCAAGACAGTGTGCTTG 180
 Db 672 ACGAGAGAGCCCAAGACACCAACCAAACTAAGTGAAGAGGAGTGAAGAGGCGGCAT 731
 181 ACGAGAGAGCCCAAGACACCAACCAAACTAAGTGAAGAGGAGTGAAGAGGCGGCAT 240
 QY 732 TCCTGAGAGCGCTGCTCGCGCTGAGAGAAAGACCCAAACCGCTTCAGAGGCTCTGG 791
 241 TCCTGAGAGCGCTGCTCGCGCTGAGAGAAAGACCCAAACCGCTTCAGAGGCTCTGG 300
 Db 792 ACGGCGAAGAGAGTTCGACCCCAAACTAAGTGAAGAGGAGTGAAGAGGCGGCAT 851
 301 ACGGCGAAGAGAGTTCGACCCCAAACTAAGTGAAGAGGAGTGAAGAGGCGGCAT 360
 QY 852 GAATCCAAATGACACAGCAGAAATGAATCTACCGAAGAAAGAAATGAATGTC 911
 351 GAATCCAAATGACACAGCAGAAATGAATCTACCGAAGAAAGAAATGAATGTC 419
 QY 912 GCCAAGAAAGATTCGAGATGAGAGAAACAGACAGTCAAGTCTCTACCAAGAGATG 971
 420 GCCAAGAAAGATTCGAGATGAGAGAAACAGACAGTCAAGTCTCTACCAAGAGATG 479
 QY 972 ATTGAGAGGATGCTGAAGAAACAGAAAGAGACAAAGAAAGAGAGAGAGAGG 1031
 480 ATTGAGAGGATGCTGAAGAAACAGAAAGAGACAAAGAAAGAGAGAGAGAGG 539
 Db 1032 AGAAGCCCAAGAGAGAGAGATGAGAAATCGATGAGAGTGAATGATGAGAGAGAA 1091
 540 AGAAGCCCAAGAGAGAGAGATGAGAAATCGATGAGAGTGAATGATGAGAGAGAA 598
 QY 1092 CAATGAAAGCCAGAGAGAAACAGTGTGTAATGTCATTAATAATGGGCAATCACTTCA 1151
 599 CAATGAAAGCCAGAGAGAAACAGTGTGTAATGTCATTAATAATGGGCAATCACTTCA 658
 Db 1152 AAGAGCTTAACAG 1211

Db 659 AAGAGCTTAACAG 718
 QY 1212 AGATGAAAG 1270
 Db 719 AGATGAAAG 778
 QY 1271 GAAAG 1330
 779 GAAAG 838
 QY 1331 GAAGCAG 1389
 Db 839 GAAGCAG 898
 QY 1390 AAGATGAG 1412
 Db 899 AAGATGAG 921

RESULT 2

AL577531/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AL577531 1201 bp mRNA linear EST 01-JUN-2003
 AL577531 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 clone CSOD1087YP20 3-PRIME, mRNA sequence.
 AL577531
 AL577531.2 GI:31315791
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 1201)
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On Feb 16, 2001 this sequence version replaced gi:12940753.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: segre@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 4268.f For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CSOD1087DH10NP1&cluster=4268.f. Contact :
 Feng liang Email: fliang@life.techn.com URL :
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID: CSOD1087DH10NP1.

FEATURES

Source

1. 1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSOD1087YP20"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="Tset strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 21.9%; Score 809.8; DB 9; Length 1201;
 Best Local Similarity 83.1%; Pred. No. 2.7e-130;
 Matches 942; Conservative 55; Mismatches 124; Indels 12; Gaps 8;

QY 2166 AAG 2225
 1144 AAG 1085
 Db 2226 AAG 2285
 1084 AAG 1025

Directionally cloned using the following adaptors:
5'-TGACACCCAGCGCTCCG-3' and
5'-GTCAGTGTGATTAATGACGCGCGCGCCCTT(15)-3'. Size selected >
1 kb for average insert length 1.87 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Iupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."

ORIGIN

Query Match	20.8%	Score 769;	DB 13;	Length 941;
Best Similarity	93.9%	Pred. No. 3.4e-123;		
Matches 821; Conservative	0;	Mismatches 51;	Indels 2;	Gaps 2;

Dp		541	AGGAGATCGGAGCAGTATACCAAGTCGATTGAGGGAAACAAAAGCTCCAAACTACAAG	600
Qy		2534	CCGGCAGCCTTCGGATCTTTCCCTGTCTCTGTGTAAGGTGACGCACAATCAAAGATATGTGG	2593
Dd		601	CCGGCAGCCTTCGGATCTTTCCCTGTCTCTGTGTAAGGTGACGCACAATCAAAGATATGTGG	660
Qy		2594	GAGAAAGGAGATGTGTTTTCATCCCCCATGACGACGACACCAATAAGAAGAACTGCT	2653
Dd		661	GAGAAAGGAGATGTGTTTTCATCCCCCATGACGACGACACCAATAAGAAGAACTGCT	720
Qy		2654	GGTTTGAGGTAGGGGCTTTTAGCCGCATCAATGAAATGGCTAATAAACCCAGATAGG-	2712
Dd		721	GGTTTGAGGTAGGGGCTTTTAGCCGCATCAATGAAATGGCTAATAAACCCAGATAGGA	780
Qy		2713	AAACAAGTCACTCTGCCCAAACTTGTGACTTGAGACCAGAGAGA	2757
Dd		781	AAACAAGCCTCACTCTGCCCAAACTTGTGACTGAGACCAAGAGAAA	825
RESULT 4				
BQ925009	LOCUS			
DEFINITION		BQ925009	941 bp mRNA linear EST 20-AUG-2002	
ACCESSION			AGENCOURT_8840364 Lupski_sciatic_nerve Homo sapiens CDNA clone	
VERSION		BQ925009	IMAGE:6205114 5', mRNA sequence.	
KEYWORDS		BQ925009.1	GI:22340040	
SOURCE		EST.		
ORGANISM		Homo sapiens (human)		
REFERENCE		Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE		1 (bases 1 to 941)		
JOURNAL		NIH-MGC http://mgc.nci.nih.gov/.		
COMMENT		National Institutes of Health, Mammalian Gene Collection (MGC)		
		Unpublished (1999)		
		Contact: Robert Strausberg, Ph.D.		
		Email: cgapds-re@mail.nih.gov		
		Tissue Procurement: Dr. James R. Lupski		
		cDNA Library Preparation: Life Technologies, Inc.		
		cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
		DNA Sequencing by: Agencourt Bioscience Corporation		
		Clone distribution: MGC clone distribution information can be		
		found through the I.M.A.G.E. Consortium/LNL at:		
		http://image.lnl.gov		
		plate: LHAM3627 row: h column: 11		
		High quality sequence stop: 617.		
FEATURES		Location/Qualifiers		
source		1..941		
		/organism="Homo sapiens"		
		/mol_type="mRNA"		
		/db_xref="taxon:9606"		
		/clone="IMAGE:6205114"		
		/sex="male"		
		/tissue_type="sciatic nerve"		
		/dev_stage="adult, 70 yr"		
		/lab_host="DHIOB"		
		/clone_id="Lupsk1_sciatic_nerve"		
		/note="Vector: pCMV-SPORT (Life Technologies); Site_1:		
		Note: Site_2: SalI; cDNA made by oligo-dT priming."		

QY	245	TTGATCTCCTGCTTAAAGAAATCAGTCCTTCCTTCCGACTTATGCTCGGGAAGAATT	304
Db	1	CTGTGCTCCTGCTTAAAGAAATCAGTCCTTCCTTCCGACTTATGCTCGGGAAGAATT	60
QY	305	TCAGACTACAGGATCATTTGGAACATTTCAAGTCAATCAAAATCCACAGGATT	364
Db	61	TCAGACTACAGGATCATTTGGAACATTTCAAGTCAATCAAAATCCACAGGATT	120
QY	365	GATGACCAACAGAGGCTCAGACATCTGATTTGCTGACTCTCCAGACATCATCTGCT	424
Db	121	GATGACCAACAGAGGCTCAGACATCTGATTTGCTGACTCTCCAGACATCATCTGCT	180
QY	425	CCCTGAACCTGAATATCACACCATGATGATTTTGAAGCTCGCAGAACTTGAAGCCA	484
Db	181	CCCTGAACCTGAATATCACACCATGATGATTTTGAAGCTCGCAGAACTTGAAGCCA	240
QY	485	AAGAGGAGAGATGCGATCGAAGCAGAAAGAAATGCGCTACAGAGATGACATGAT	544
Db	241	AAGAGGAGAGATGCGATCGAAGCAGAAAGAAATGCGCTACAGAGATGACATGAT	300
QY	545	GAAGGAGAGGCGCCCTCGGGAACGGCGCGCGCGACCCGACAGGAACGGCTCGGCGAAG	604
Db	301	GAAGGAGAGGCGCCCTCGGGAACGGCGCGCGCGACCCGACAGGAACGGCTCGGCGAAG	360
QY	605	CAGAGGAGAAATCTTTGGGACAGGTGACCCGACCAAGTGAAGTGAATGCCAGAACGT	664
Db	361	CAGAGGAGAAATCTTTGGGACAGGTGACCCGACCAAGTGAAGTGAATGCCAGAACGT	420
QY	665	GTGCGCTGACGAGAGGSCAAAGAACCAACCAACAACCTCAAGTGAAGAGGATGATGAG	724
Db	421	GTGCGCTGACGAGAGGSCAAAGAACCAACCAACAACCTCAAGTGAAGAGGATGATGAG	480
QY	725	GCCGCAATTCCTGGAGCGCGCTTGAGCTCGGCGTGAAGAAAGACGCAAAAACGCTTCAGAG	784
Db	481	GCCGCAATTCCTGGAGCGCGCTTGAGCTCGGCGTGAAGAAAGACGCAAAAACGCTTCAGAG	540
QY	785	GCTCTGGAGCGGAGAGAGGATTGACCCCAACATTAACAGATGCAAGTCTGTGCTCCCA	844
Db	541	GCTCTGGAGCGGAGAGAGGATTGACCCCAACATTAACAGATGCAAGTCTGTGCTCCCA	600
QY	845	AGCAGAGAGATGCAAAATGACACACGACGAAATGAAATCTACCGAGAGGAAGAAAAAGT	904
Db	601	AGCAGAGAGATGCAAAATGACACACGACGAAATGAAATCTACCGAGAGGAAGAAAAAGT	660
QY	905	GAAGGTCGCGCAAGAAAGATACAGAGATGAGAGAAACAGAAACAGTCAACCAAGTCTACAG	964
Db	661	GAAGGTCGCGCAAGAAAGATACAGAGATGAGAGAAACAGAAACAGTCAACCAAGTCTACAG	720
QY	965	AAGATGATT-GGAGGATGCTGTGAAGAAAACAAGAAAGAAAGACAGAGAAA-GGAGAGG	1022
Db	721	AAGATGATTGGAGGATGCTGTGAAGAAAACAAGAAAGAAAGACAGAGAAAAGGAGGAGG	780
QY	1023	AGGAAGAGAGAAAGCCAAAGCGAGAGGAGCATTTGAGAAAAATCAGGTGAAGTATGTTGG	1082
Db	781	AGAAATAGAGAAAGCCCAAGCCGAGGAGGATGAGGAGAAAAATCGATCCAAGATGAAA	840
QY	1083	AAGAGAAAACATGAAAGCCAGAGAGAAACGT 1116	
Db	841	GATTAAAAAGACCAAGAGACCCCAAGAGAAAGT 874	

RESULT 5
BU151076 846 bp mRNA linear EST 03-SEP-2002
LOCUS AGENCOURT_8586111.Lupski_sym pathetic_trunk Homo sapiens cDNA clone
DEFINITION IMAGE:6195568 5', mRNA sequence.
ACCESSION BU151076
VERSION BU151076.1 GI:22664608
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 846)
NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgarbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLNL13602 row: 3 column: 17
High quality sequence stop: 686.
Location/Qualifiers
1. 846
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6195568"
/sex="male"
/tissue_type="sympathetic trunk"
/dev_stage="adult, 16 yr"
/lab_host="DH10B"
/clone_lib="Lupski_sym pathetic_trunk"
/note="Vector: PCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TGACCCACGCGTCCG-3' and
5'-GACTAGTCTAGATCGGAGCGCCGCTT(5'-3', Size selected >
1 kb for average insert length 1.9 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine); available through Life
Technologies."

ORIGIN
Query Match 20.8%; Score 768.2; DB 13; Length 846;
Best Local Similarity 96.3%; Pred. No. 4.8e-123;
Matches 808; Conservative 0; Mismatches 28; Indels 3; Gaps 2;
QY 2437 ATGACCCATCAAGAGCAATAGTCTCCAGATTGACGACGACTGAGCAGTATACCG 2496
DB 1 ATGACCCATCAAGAGCAATAGTCTCCAGATTGACGACGACTGAGCAGTATACCG 60
QY 2497 TGCAATTGAGGGAACAAAAGCGCAAACTCAAAAGCGGAGCGCTCGATCTTCCTGT 2556
DB 61 TGCAATTGAGGGAACAAAAGCGCAAACTCAAAAGCGGAGCGCTCGATCTTCCTGT 120
QY 2557 TCTGTGTTGAAGGTGTACGCAATCAAGATGTGTGGAGAAAGGAATGTGTTTCATC 2616
DB 121 TCTGTGTTGAAGGTGTACGCAATCAAGATGTGTGGAGAAAGGAATGTGTTTCATC 180
QY 2617 CCCCATGCGAGCGAGCAGCAATTAAGAAATGCTGGCTTGAAGTAGGGGTTTCTAG 2676
DB 181 CCCCATGCGAGCGAGCAGCAATTAAGAAATGCTGGCTTGAAGTAGGGGTTTCTAG 240
QY 2677 CCCCATCAATGATGGCTAACTAAACCCAGATGAGAAACAAGTCACTGCTCCCAACC 2736

DB 241 CCCCATCAATGATGGCTAACTAAACCCAGATGAGAAACAAGTCACTGCTCCCAACC 300
QY 2737 TTCTGACTTGAAGACAGAGACGATATCCAGCAGCGGAACCTCTGGGAAACGATCTGT 2796
DB 301 TTCTGACTTGAAGACAGAGACGATATCCAGCAGCGGAACCTCTGGGAAACGATCTGT 360
QY 2797 GGATTAAGGTCACTTCCCCCACTAAGTTTGAACAGCTTCCAGAAAGAACCCAGCTCAAG 2856
DB 361 GGATTAAGGTCACTTCCCCCACTAAGTTTGAACAGCTTCCAGAAAGAACCCAGCTCAAG 420
QY 2857 ACCGAGACAGAGCTCACTTGTAGAGGCGTATATTCGCTGTTTATTTATTTATTTAT 2916
DB 421 ACCGAGACAGAGCTCACTTGTAGAGGCGTATATTCGCTGTTTATTTATTTATTTAT 480
QY 2917 ACTAAATTGGGTTCACTTATCTTTTATTTTCAATATCCAGTAAACCCATGATATATTC 2976
DB 481 ACTAAATTGGGTTCACTTATCTTTTATTTTCAATATCCAGTAAACCCATGATATATTC 540
QY 2977 ACTAATTATTAATCAAGCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 3036
DB 541 ACTAATTATTAATCAAGCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 600
QY 3037 AGCCTGTTTCAAGAAACCCATGCTGTGAATAGAGCTTTTCTAATCATCATCAAC 3096
DB 601 AGCCTGTTTCAAGAAACCCATGCTGTGAATAGAGCTTTTCTAATCATCATCAAC 659
QY 3097 TCTGTATCTGAGCGATATACCAACCATGCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 3156
DB 660 TCTGTATCTGAGCGATATACCAACCATGCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 719
QY 3157 TTGCTGCGGAATGTGTGAGTATCTAGAAAATGAACCGTAGTTTGTGTTT- -AA 3214
DB 720 TTGCTGCGGAATGTGTGAGTATCTAGAAAATGAACCGTAGTTTGTGTTT- -AA 779
QY 3215 ATACAGAGTCATGTTTCTGACCTTATTAATTAACATGAGAAATTAATCTAGT 3273
DB 780 TACAGAAATGATGTTGTTTCTGACCTTATTAATTAACATGAGAAATTAAT 838

RESULT 6
CD359921 881 bp mRNA linear EST 29-MAY-2003
LOCUS AGENCOURT_14275922.NIH.MGC.180.Homo.sapiens.cDNA.clone
DEFINITION IMAGE:30358440 5', mRNA sequence.
ACCESSION CD359921
VERSION CD359921.1 GI:31131332
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 881)
NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgarbs-remail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: NDAM460 row: 5 column: 01
High quality sequence stop: 622.
Location/Qualifiers
1. 881
/organism="Homo sapiens"

ORIGIN

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30388440"
/lab_host="DH10B-Ton A (T1 and T5 phage resistancees)"
/clone_lib="NIH MGC 180"
/Note="Organ: Testis; Vector: pCMV-Sport6.1; Site: 1; NotI; Site 2: EcoRV (destroyed); Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.68 kb. Library was constructed by (Invitrogen). Note: this is a NIH-MGC Library."

Query Match      20.7%; Score 768; DB 14; Length 881;
Best Local Similarity 99.2%; Pred. No. 5.1e-123;
Matches 771; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 CCTAGGGAATGACAGGCACTCTCCACAGGAGGCTGCATCCACCTTGCTGGGGTGTCTC 61
Db 1 CCTAGGGAATGACAGGCACTCTCCACAGGAGGCTGCATCCACCTTGCTGGGGTGTCTC 60
QY 62 ATTGGCTGCTTATGAAAAACGACAGGCAATGCAACACCGCTCCCGACTGTAAAC 121
Db 61 ATTGGCTGCTTATGAAAAACGACAGGCAATGCAACACCGCTCCCGACTGTAAAC 120
QY 122 ATAGGGGATATGTGTCACTTAGCATGGACTTCTGGGAGGGCCAAAGAGGCGGTCTG 181
Db 121 ATAGGGGATATGTGTCACTTAGCATGGACTTCTGGGAGGGCCAAAGAGGCGGTCTG 180
QY 182 GAGTTTATTAATGATGAGCAGTGTGATTCGGCTGCTGCCTGCGCGCTGTGCTCTC 241
Db 181 GAGTTTATTAATGATGAGCAGTGTGATTCGGCTGCTGCCTGCGCGCTGTGCTCTC 240
QY 242 TGCGTGTGCTCTGTTAAAGAAATCACTCTTCTTCGACTTGAATCTCGGGAAGA 301
Db 241 TGCGTGTGCTCTGTTAAAGAAATCACTCTTCTTCGACTTGAATCTCGGGAAGA 300
QY 302 GTTTCAGCTCAAGGTATCATTTGAAACATTTCAAGATCATCAATCAATTCACAGGG 361
Db 301 GTTTCAGCTCAAGGTATCATTTGAAACATTTCAAGATCATCAATCAATTCACAGGG 360
QY 362 ATTGGTGACCAACGAAAGGCTCAGACATCTGATTTCTGACTGCTGCACATCATCTGG 421
Db 361 ATTGGTGACCAACGAAAGGCTCAGACATCTGATTTCTGACTGCTGCACATCATCTGG 420
QY 422 TCTCCCTGACCTGAAATCAACACATGATGATTTGAGGCTGCGAGAACTTAGAGG 481
Db 421 TCTCCCTGACCTGAAATCAACACATGATGATTTGAGGCTGCGAGAACTTAGAGG 480
QY 482 CAAAAGAGGAGAGATGCTCGAAGCAAAAGATGCTTACCAAGAAATGACAT 541
Db 481 CAAAAGAGGAGAGATGCTCGAAGCAAAAGATGCTTACCAAGAAATGACAT 540
QY 542 GATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 601
Db 541 GATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
QY 602 AAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 661
Db 601 AAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 662 AGTGTGCTGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 721
Db 661 AGTGTGCTGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 722 GAGGCGGCAATCTCTGAGAGGCTCTGCTGCGCGGTGAGAGAGAGAGAGAGAGAGAGAG 778
Db 721 GAGGCGGCAATCTCTGAGAGGCTCTGCTGCGCGGTGAGAGAGAGAGAGAGAGAGAGAG 777

```

RESULT 7
CD643635 827 bp mRNA linear EST 17-JUN-2003
LOCUS CD643635

```

DEFINITION
AGENCOURT_14553720 N1A Human H1 Embryonic Stem Cell cDNA Library
(Long) Homo sapiens cDNA clone IMAGE:30425920 5', mRNA sequence.
CD643635
CD643635.1 GI:31814712
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 1 to 827)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics / NIH
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabs-remail.nih.gov
Tissue Procurement: Irene Gius and Mahendra Rao, NIA
CDNA Library Preparation: Yulan Piao and Minoru Ko
DNA Sequencing by: Agencourt Bioscience Consortium (LNL)
clone distribution: MGC c
http://image.llnl.gov
plate: NDM510 row: 0 column: 17
High quality sequence stop: 662.
Location/Qualifiers
1..827
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30425920"
/tissue_type="Embryonic Stem cells"
/clone_lib="N1A Human H1 Embryonic Stem Cell cDNA Library (Long)"
/Note="Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI; This is a long-transcript enriched cDNA library (genome Res. 11: 1553-1558 (2001)). [PMID: 11544199] From WA01 cell line. Undifferentiated human ES cell line WA01/H1 was obtained from Wicell Research Institute, Inc. Madision, WI, cultured according to their instructions, on MEf feeders. They formed round colonies with defined edges and were positive for alkaline phosphatase, SSEA-4, OCT3, OCT4, REX1, UTR, TERT, SOX2, CX43 and CX45. They are negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLK3, SSEA-1, TUBB3, NES, GFAP, and BOMBS. When confluent (18-10 days after plating), the ES cells from 4 X 6cm dishes were treated with 1 mg/ml collagenase, type IV (Invitrogen/Gibco) for 5-10 min and gently scraped off with 5 ml pipette. RNA was purified with RNeasy Reagent from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558 (2001). [PMID: 11544199] Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTAGTCTGATCGAGCGCGCCCTTTTCTTTTCTTTT-3'] from 3.46 of total RNA, treated with 14 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Loner-linker LB-salI, purified by phenol/chloroform extraction, and separated from free linkers by Centricon-100 column. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex tag polymerase (Takara) with a primer Sal4-S for 25 cycles. The products were purified by phenol/chloroform extraction and Centricon-100 column. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-Sport6 plasmid vector. The average insert size is about 3.6kb."

```

FEATURES

```

source
Query Match      20.4%; Score 753.8; DB 14; Length 827;
Best Local Similarity 97.5%; Pred. No. 1.5e-120;

```

ORIGIN

Query Match 20.4%; Score 753.8; DB 14; Length 827;
Best Local Similarity 97.5%; Pred. No. 1.5e-120;

Matches 769; Conservative 0; Mismatches 19; Indels 1; Gaps 1;

QY 238 TCTTGGCTGCTGCTCTGCTTTAAAGAAATGACGCTCTTCCGACTTAGTCTCGGGA 297

Db 9 TCCCTGGCTGCTGCTCTGCTTTAAAGAAATGACGCTCTTCCGACTTAGTCTCGGGA 68

QY 298 AGAAGTTGACATACAGATATCATTTGGAACATTTCAAGATCATCAATCAATTCAC 357

Db 69 AGAAGTTGACATACAGATATCATTTGGAACATTTCAAGATCATCAATCAATTCAC 128

QY 358 AGGAGTTGTTGACCAACAGAGGCTCAGACATCTGATTTGCTGACCTGTCAGACATCAT 417

Db 129 AGGAGTTGTTGACCAACAGAGGCTCAGACATCTGATTTGCTGACCTGTCAGACATCAT 188

QY 418 CTGGCTCTCCCTGAAACCTGAAATCAACACATGATGATTTTGAAGCTGCGAGAACTTNG 477

Db 189 CTGGCTCTCCCTGAAACCTGAAATCAACACATGATGATTTTGAAGCTGCGAGAACTTNG 248

QY 478 AAGGCAAAAGAGGAGAGATGCGACTCGAAGCAGAAAGAAATCGCTTACAGAGATGA 537

Db 249 AAGGCAAAAGAGGAGAGATGCGACTCGAAGCAGAAAGAAATCGCTTACAGAGATGA 308

QY 538 CGATGATGAAAGAGAGGAG 597

Db 309 CGATGATGAAAG 368

QY 598 GGAAG 657

Db 369 GGAAG 428

QY 658 GAAAG 717

Db 429 GAAAG 488

QY 718 TGATGAG 777

Db 489 TGATGAG 548

QY 778 TCAG 837

Db 549 TCAG 608

QY 838 GCTCCCAAG 897

Db 609 GCTCCCAAG 668

QY 898 AAAAAGTGAAGTGCAG 957

Db 669 AAAAAGTGAAGTGCAG 728

QY 958 CTACCGAG 1017

Db 729 CTACCGAG 787

QY 1018 GGAG 1066

Db 788 GGAG 796

RESULT 8

LOCUS AL700988

DEFINITION DKF2686D12120.r1.686 (synonym: h1c3) Homo sapiens cDNA clone

ACCESSION AL700988

VERSION AL700988

KEYWORDS EST.19621521

SOURCE EST. Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 749)

AUTHORS Ootemwaelder, B., Obermaier, B., Mewes, W., Mewes, H.W., Weill, B. and

TITLE Wiemann, S., Ootemwaelder, B., Obermaier, B., Mewes, H.W., Weill, B. and

JOURNAL Wiemann, S.

COMMENT Unpublished (2001)

CONTACT: MIPS

MIPS Ingolstaedter Landstr. 1, D-85764 Neuberg, Germany

This is the 5' sequence of the clone insert

Research from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de; sequenced by MedGenomix (Martinriedel/Germany) within the cDNA sequencing consortium of the German Genome Project. No si sequence available.

This clone (DKF2686D12120) is available at the RZPD in Berlin. Please contact the RZPD: Resourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source

1..749

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="DKF2686D12120"

/dev_stage="adult"

/lab_host="DH10B"

/clone_id="686 (synonym: h1c3)"

/note="Vector: pTriplex2; site_1: sf1a; site_2: sf1b; cDNA-collection"

ORIGIN

Query Match 20.2%; Score 747.4; DB 9; Length 749;

Best Local Similarity 99.9%; Pred. No. 26-119; Indels 0; Gaps 0;

Matches 749; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 201 AGTGTATTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTAA 260

Db 1 AGTGTATTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTAA 60

QY 261 AGAATAGAGCTTCTCTTCCGACTTATCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAT 320

Db 61 AGAATAGAGCTTCTCTTCCGACTTATCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAT 120

QY 321 CATTGGAACATTTCAAGATCATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 380

Db 121 CATTGGAACATTTCAAGATCATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 180

QY 381 GCTCAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 440

Db 181 GCTCAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240

QY 441 AACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 500

Db 241 AACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300

QY 501 GACTCGAAGCAGAAAGATTCGCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 560

Db 301 GACTCGAAGCAGAAAGATTCGCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360

QY 561 GGAAG 620

Db 361 GGAAG 420

QY 621 TGGAG 680

Db 421 TGGAG 480

QY 681 CCAAG 740

Db 481 CCAAG 540

QY 741 GCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 800

Db 541 GCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600

QY 801 AGGAGTTGACCCACATATAAGATGAGTGTCTGCTCCCAAGCAGAGATGCAAA 860
 Db 601 AGGAGTTGACCCACATATAAGATGAGTGTCTGCTCCCAAGCAGAGATGCAAA 860
 QY 861 ATGACACACGAGAAATTAATCTACCGGAGAGAGAAAGAAAGTGAATTCGCGCAGAAA 920
 Db 661 ATGACACACGAGAAATTAATCTACCGGAGAGAGAAAGAAAGTGAATTCGCGCAGAAA 920
 QY 921 GATACGAGATAGAGAAACAGAAACAGTC 949
 Db 721 GATACGAGATAGAGAAACAGAAACAGTC 949

RESULT 9
 AUI32922 812 bp mRNA linear EST 01-AUG-2002
 LOCUS AUI32922 NT2RP4 Homo sapiens cDNA clone NT2RP4000885 5', mRNA
 DEFINITION
 AUI32922.1 GI:10993461
 VERSION AUI32922.1
 KEYWORDS Homo sapiens (human)
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 812)
 Ota,T., Sugiyama,T., Ishii,S., Suzuki,Y., Saito,K., Yamamoto,J., Nishikawa,T., Nakamura,Y., Nagai,T., Sugano,S., Masuno,Y. and Isegai,T.
 HRI human cDNA project (Ota,T., Sugiyama,T., Ishii,S., Suzuki,Y., Saito,K., Yamamoto,J., Nishikawa,T., Nakamura,Y., Nagai,T., Sugano,S., Masuno,Y., Isegai,T.)
 Unpublished (2000)
 Contact: Takao Isegai
 Genomics Laboratory
 Helix Research Institute
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: genomics@hri.co.jp
 HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix Research Institute; cDNA library construction; Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.
 Location/Qualifiers
 1..812
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="NT2RP4000885"
 /cell_type="teratocarcinoma"
 /clone_1b="NT2"
 /clone_1b="NT2RP4"
 /note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction"

ORIGIN
 Query Match 19.9%; Score 737.8; DB 9; Length 812;
 Best Local Similarity 97.9%; Fred. No. 8.9e-118;
 Matches 799; Conservative 0; Mismatches 10; Indels 7; Gaps 5;

QY 1910 ACAGAAAGTTAAGTCGAGATGAGATTCATGACCCACAACTTAACATCTAGAAAT 1969
 Db 1 ACAGAAAGTTAAGTCGAGATGAGATTCATGACCCACAACTTAACATCTAGAAAT 60

QY 1970 ACTTTCAGCCGCCCTGAGAGAGAGCCGAGGCTGAGACCAAGAGGCTGAGGCGCCGCC 2029
 Db 61 ACTTTCAGCCGCCCTGAGAGAGAGCCGAGGCTGAGACCAAGAGGCTGAGGCGCCGCC 120

QY 2030 CAGGTGGAAGCCGCGCAAAAGGCTGAGAGAGCTTGTGTGCTGCGCGGAGAGACCGAGAGC 2089
 Db 121 CAGGTGGAAGCCGCGCAAAAGGCTGAGAGAGCTTGTGTGCTGCGCGGAGAGACCGAGAGC 180

QY 2090 GAAAGTTGAGAAAGCTCAACAGAGAGCGCAGAGCGGCTTTGGAGCTGAGAACTC 2149
 Db 181 GAAAGTTGAGAAAGCTCAACAGAGAGCGCAGAGCGGCTTTGGAGCTGAGAACTC 240

QY 2150 AAGAAAAAGAGGAG 2209
 Db 241 AAGAAAAAGAGGAG 300

QY 2210 GAGGAAGCCGATCGAAAACTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2269
 Db 301 GAGGAAGCCGATCGAAAACTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360

QY 2270 GAAAGGCGAG 2329
 Db 361 GAAAGGCGAG 420

QY 2330 GATGACAAAGAAACCATTCAGAGTGTCTTCACTCTTAAAGTTTCAATCTCAAGATGAAAG 2389
 Db 421 GATGACAAAGAAACCATTCAGAGTGTCTTCACTCTTAAAGTTTCAATCTCAAGATGAAAG 480

QY 2390 CGAGCAGAAATTTTGAATPAAGTCTGTGAGAAAGAGAGAGAGAGAGAGAGAGAGAG 2449
 Db 481 CGAGCAGAAATTTTGAATPAAGTCTGTGAGAAAGAGAGAGAGAGAGAGAGAGAGAG 537

QY 2450 GCAGCAATAGTCTCCAGATTGACAGACAGACTGAGACAGATACCACTGCAATTGAGGGA 2509
 Db 538 GCAGCAATAGTCTCCAGATTGACAGACAGACTGAGACAGATACCACTGCAATTGAGGGA 597

QY 2510 ACAAAGACCGCAAAACCTTCAAG--CCGCAAGCTGGATCTCTGCTGCTGAGAG 2568
 Db 598 ACAAAGACCGCAAAACCTTCAAG--CCGCAAGCTGGATCTCTGCTGCTGAGAG 657

QY 2569 TGTAGCAACATCAAGAGATGTGGAGAAAGGAAATGTGTTTCACTCCAC--TGCAG 2627
 Db 658 TGTAGCAACATCAAGAGATGTGGAGAAAGGAAATGTGTTTCACTCCAC--TGCAG 717

QY 2628 CAGGACACCAATTAAGAAAC--TGCTGCTTGAAGTAGGGTTTGAAGCGATCAAT 2686
 Db 718 CAGGACACCAATTAAGAAAC--TGCTGCTTGAAGTAGGGTTTGAAGCGATCAAT 776

QY 2687 GAATGCTACTTAAACCCAGATGAGAAACCACTCA 2722
 Db 777 GAATGCTACTTAAACCCAGATGAGAAACCACTCA 812

RESULT 10
 CD656598 823 bp mRNA linear EST 18-JUN-2003
 LOCUS AGENCOURT 14540446 NIA Human HI Embryonic Stem Cell cDNA Library
 DEFINITION (Long) Homo sapiens cDNA clone IMAGE:30419432 5', mRNA sequence.
 CD656598
 AGENCOURT 14540446 NIA Human HI Embryonic Stem Cell cDNA Library
 VERSION CD656598.1 GI:31898079
 KEYWORDS EST
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 823)
 NIH-MGC http://img.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC).
 Unpublished (1999)
 Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 31d9.31 Rm10A07 Bethesda, MD 20892
 Email: cgsabrs-remail.nih.gov
 Tissue Procurement: Irene Givis and Mahendra Rao, NIA
 cDNA Library Preparation: Yulan Piao and Minoru Ko
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNU)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC
 Clone distribution information can be found through the I.M.A.G.E. Consortium/ILNU at:
 http://image.llnl.gov

Plate: NDAM494 row: a column: 09
High quality sequence stop: 695.
Location/Qualifiers
1..823
source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30419432"
/tissue_type="Embryonic Stem cells"
/cell_line="WA01"
/lab_host="DH10B (T1 Phase-resistant)"
/clone_1lb="NIA Human H1 Embryonic Stem Cell cDNA Library (long)"
/note="Vector: pCMV-Sport6, Site_1: NotI, Site_2: SalI; This is a long-transcript enriched cDNA library (Genome Res. 11: 1553-1558 (2001)). [PMID: 11541991] from WA01 cell line. Undifferentiated human ES cell line WA01/H1 was obtained from WiCell Research Institute, Inc., Madison, WI, cultured according to their instructions, on MEK feeders. They formed round colonies with defined edges and were positive for alkaline phosphatase, SSEA-4, OCT3, OCT4, REX1, UTR, TERT, SOX2, CX43 and CX45. They are negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLT3, SSEA-1, TUBB3, NES, GFAP, and EOMES. When confluent (18-10 days after plating), the ES cells from 4 X 6cm dishes were treated with 1 mg/ml collagenase, type IV (Invitrogen/GIBCO) for 5-10 min and gently scraped off with 5 ml pipette. RNA was purified with TRIzol Reagent from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558 (2001). [PMID:11541991] Double-stranded cDNAs were synthesized with an Oligo(dT) primer (Invitrogen: 5'-pGAGTACTCTAGATCGGAGCGCGCCCTTTTCTTTT-3') from 3.4g of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Loxe-linker IL-Sal4, purified by phenol/chloroform extraction, and separated from free linkers by Centricon-100 column. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S for 25 cycles. The products were purified by phenol/chloroform extraction and Centricon-100 column. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-Sport6 plasmid vector. The average insert size is about 3.6kb."

ORIGIN

Query Match 19.8%; Score 732; DB 14; Length 823;
Best Local Similarity 97.5%; Pred. No. 9e-117;
Matches 752; Conservative 0; Mismatches 18; Indels 1; Gaps 1;

QY 260 AAGAAATCACTCTCTCTTCCGACTTGTCTCGGGAAGAGTTTCACTACAAAGTA 319
Db 12 AAGAAATCACTCTCTCTTCCGACTTGTCTCGGGAAGAGTTTCACTACAAAGTA 71
QY 320 TCATTGGACATTTCAGATCATCAATCAATTCCACAGGATGTGACCAACAGAA 379
Db 72 TCATTGGACATTTCAGATCATCAATCAATTCCACAGGATGTGACCAACAGAA 131
QY 380 GGGTCAGACATCTGATTGCTGACCTGTCCAGACATCTGTGCTTCCCTGAACCTGAAT 439
Db 132 GGGTCAGACATCTGATTGCTGACCTGTCCAGACATCTGTGCTTCCCTGAACCTGAAT 191
QY 440 CACACATGATGATTTTGGAGCGTCGACAGAACTTAGAAGGCAAAAGAGAGGAGATG 499
Db 192 CACACATGATGATTTTGGAGCGTCGACAGAACTTAGAAGGCAAAAGAGAGGAGATG 251
QY 500 CGACTGGAAGCAAGAAAGATGCTTACAGAGAAATGACGATGATGAAGAGAGGAGCC 559
Db 252 CGAGCCGGAACAGAAAGATGCTTACAGAGAAATGACGATGATGAAGAGAGGAGCC 311
QY 560 CGGAAACGGCGCCCGAGCCCGACAGGAACGGTCGGCAGAGAACAGAGAGAGATCC 619
Db 312 CGGAAACGGCGCCCGAGCCCGACAGGAACGGTCGGCAGAGAACAGAGAGAGATCC 371

QY 620 TTGGGACAGGTGACCGACCAAGGTGAGGTGATGCCCAGAAACAGTGTGCTGACGAGAG 679
Db 372 TTGGGACAGGTGACCGACCAAGGTGAGGTGATGCCCAGAAACAGTGTGCTGACGAGAG 431
QY 680 GCCAAGCAACACCAACAAACATCTCAGTGGAAAGGGATGATGAGCCGCAATTCCTGGAG 729
Db 432 GCCAAGCAACACCAACAAACATCTCAGTGGAAAGGGATGATGAGCCGCAATTCCTGGAG 491
QY 740 CGCTGGCTCGCGGTGAGAAAGACGCCAAAACGCCCTTCAGAGGCTTGGAGCGGAG 799
Db 492 CGCTGGCTCGCGGTGAGAAAGACGCCAAAACGCCCTTCAGAGGCTTGGAGCGGAG 551
QY 800 AAGAGTTGACCCCAACATATACAGATGCAAGTCTGTGCTCCCAAGCAAGAAAGCA 859
Db 552 AAGAGTTGACCCCAACATATACAGATGCAAGTCTGTGCTCCCAAGCAAGAAAGCA 611
QY 860 AATGACACACAGAAATGAACTACCGAGAGAAAGAAAGAAAGTGAAGTCCCAAGAA 919
Db 612 AATGACACACAGAAATGAACTACCGAGAGAAAGAAAGAAAGTGAAGTCCCAAGAA 670
QY 920 AGATGACAGTATGAGGAAACAGAAACAGTCCCAAGTCTTACCAAGAAATGATGGAG 979
Db 671 AGATGACAGTATGAGGAAACAGAAACAGTCCCAAGTCTTACCAAGAAATGATGGAG 730
QY 980 GATGCTGAAGAAACAGAAAGAGACAAAGAAAGAGAGAGAGAGAGAG 1030
Db 731 GATGCTGAAGAAACAGAAAGAGACAAAGAAAGAGAGAGAGAGAGAGAGAG 781

RESULT 11
BO010124/c
LOCUS
DEFINITION
IMAGE:5835886 3', mRNA sequence.

ACCESSION
BO010124
BO010124.1 GI:19735025
EST.
KEYWORDS
Homo sapiens (human)
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 748)
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Dr. Jose Mercende
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/ILM at: <http://image.llnl.gov>
The following repetitive elements were found in this cDNA
sequence: 1-21, >At rich<low complexity (matched complement)
Seq primer: M13 FORWARD
PolyA-yes.

FEATURES

Location/Qualifiers
1..748
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5835886"
/tissue_type="Chondrosarcoma"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/clone_id="NCI-CCAP_ED0"
/note="Organ: Left Pubic Bone; Vector: pTTT3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CCAP ED0 is a cDNA library containing the following tissue(s): Chondrosarcoma cell line C85. The

library was constructed according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT733-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dr)18 tail. The sequence tag for this library is GCTCAAGCT.
TAG TISSUE=chondrosarcoma
TAG LIB=UI-H-ED0
TAG_SEQ=GCTCAAGCT"

ORIGIN

Query Match 19.6%; Score 725.4; DB 12; Length 748;
Best Local Similarity 99.1%; Pred. No. 1.3e-115;
Matches 740; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

2611 TTCAATCCCACTGAGCAGGACACCAATTAAGAACTGCTGCTGAAGTAGGGGT 2670
748 TTCAATCCCACTGAGCAGGACACCAATTAAGAACTGCTGCTGAAGTAGGGGT 690
2671 TTCAAGCGCATCAATGAATGGCTAACTAAACCCAGATGGAACAAGTCACCTGCTCC 2730
689 TTCAAGCGCATCAATGAATGGCTAACTAAACCCAGATGGAACAAGTCACCTGCTCC 630
2731 CAAACCTTCTGACTTGAGACCAAGAGACGATATCCAGAAACCGAATCTCTGGAAAAACA 2790
629 CAAACCTTCTGACTTGAGACCAAGAGACGATATCCAGAAACCGAATCTCTGGAAAAACA 570
2791 ATCTGTGATAGGTCATCTCCCACTTAAGTTGAGACAGTTCCAGAAAGAACCAAG 2850
569 ATCTGTGATAGGTCATCTCCCACTTAAGTTGAGACAGTTCCAGAAAGAACCAAG 510
2851 CTCAAGACGACGAGACGAGTCAGTTGTAGAGGGCTTAATGCTCTGTTTGTATATGT 2910
509 CTCAAGACGACGAGACGAGTCAGTTGTAGAGGGCTTAATGCTCTGTTTGTATATGT 450
2911 TGATTTACTAATTTGGTTCATTTATCTTTTCAATATCCAGTAACCCATGTAT 2970
449 TGATTTACTAATTTGGTTCATTTATCTTTTCAATATCCAGTAACCCATGTAT 390
2971 ATTATCACTAATTTAATTAATACAGTCTAGAGATGTCATGGTAAAGTACGCTTGG 3030
389 ATTATCACTAATTTAATTAATACAGTCTAGAGATGTCATGGTAAAGTACGCTTGG 330
3031 CACAGGAGCCTGTTTCTAAGAAACCCATGCTGTGAATAGAGACTTTTCTACTGATCAT 3090
329 CACAGGAGCCTGTTTCTAAGAAACCCATGCTGTGAATAGAGACTTTTCTACTGATCAT 270
3091 CATATCTGTATCTGAGAGAGTATACCAACCACTGTAAGTCAACAGAGATCCAGT 3150
269 CATATCTGTATCTGAGAGAGTATACCAACCACTGTAAGTCAACAGAGATCCAGT 210
3151 TTAATATGCTGCGGAATGTGTGAGATCTAGAAAAATGAACCGTAGTTTGTGTTT 3210
209 TTAATATGCTGCGGAATGTGTGAGATCTAGAAAAATGAACCGTAGTTTGTGTTT 150
3211 TTAATATGAGAGTATGTTGTTTCTGCACTTATATATAAGATGGAAGAAATTAATCTT 3270
149 TTAATATGAGAGTATGTTGTTTCTGCACTTATATATAAGATGGAAGAAATTAATCTT 90
3271 AGTAGGCAATTTGAACCTTTTGAAGATTAACCATTTTCAGATTGAAAAATCTGCCATTA 3330
89 AGTAGGCAATTTGAACCTTTTGAAGATTAACCATTTTCAGATTGAAAAATCTGCCATTA 30
3331 TGGTGTCTTAAAAAAGAAAAAGA 3357
29 TGGTGTCTTAAAAAAGAAAAAGA 3

RESULT 12

AT198679/c
LOCUS AT198679 775 bp mRNA linear EST 06-JUL-1999
DEFINITION w91c12.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:2348470 3' similar to SW:CALD_HUMAN Q05682 CALDESMON ; mRNA sequence.
ACCESSION AT198679
VERSION AT198679.1 GI:5364151
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 775)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapb-remail.nih.gov
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco
High quality sequence stop: 465.
Location/Qualifiers
1..775
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2348470"
/lab_host="DH10B"
/clone_lib="Soares_NFL_T_GBC_S1"
/note="Organ: pooled; Vector: pT733-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung Nbl19W, testis NHT, and B-cell NCI CGAP GCBI) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of 1.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 728096-733399. Subtraction by Bento Soares and M. Fatima Bonaldi."

ORIGIN

Query Match 19.4%; Score 719.6; DB 9; Length 775;
Best Local Similarity 97.5%; Pred. No. 1.3e-114;
Matches 750; Conservative 0; Mismatches 17; Indels 2; Gaps 2;

2573 CGCAACATCAAGAGTATGTGGAGAAAGGAATGTGTTTCATCTCCCACTGACGAGGC 2632
769 CGCAACATCAAGAGTATGTGGAGAAAGGAATGTGTTTCATCTCCCACTGACGAGGC 711
2633 ACACC-AAATAGAGAACTGCTGCTTGAAGTAGGGGTTTCTAGCCGATCAATGATG 2691
710 ACACCAAATAGAGAACTGCTGCTTGAAGTAGGGGTTTCTAGCCGATCAATGATG 651
2692 GCTAATCAAAACCCAGATGAAACAAGTCACTGCTCCAAACCTTCTGACTGAGACC 2751
650 NCTAATCAAGAGCCAGAAAGAAACAAGTCACTGCTCCAAACCTTCTGACTGAGACC 591
2752 AGGAGAGCTATCAGCAAGCGGAACCTCTGGGAAAAAGCAATCTGTGATAGTCACTTC 2811
590 AGGAGAGCTATCAGCAAGCGGAACCTCTGGGAAAAAGCAATCTGTGATAGTCACTTC 531
2812 CCCCACTAAGTTTGAAGATGCCAGAAAGAACCCAGAGTCAAGAGGAGAGAGAGTC 2871
530 CCNCACTAAGTTTGAAGATGCCAGAAAGAACCCAGAGTCAAGAGGAGAGAGAGTC 471
2872 AGTTGTAGAGGCTAATTCGCTGCTGTTTGTATTTATGTTGATTTACTAAATGGCTCA 2931
470 AGTTGTAGAGGCTAATTCGCTGCTGTTTGTATTTATGTTGATTTACTAAATGGCTCA 411

Qy	2932	TTATCTTTTATTTTTCATATCCGAGTAAACCCAGTATATATACATATATTAAAT	2991
Db	410	TTATCTTTTATTTTTCATATCCGAGTAAACCCAGTATATATACATATATTAAAT	351
Qy	2992	CACAGCTAGAGATCTTCATAGTAAABAATACGCGCTTGACAGAGCGCTGTTCTAAG	3051
Db	350	CACAGCTAGAGATCTTCATAGTAAABAATACGCGCTTGACAGAGCGCTGTTCTAAG	291
Qy	3052	AAACCCATGCTGTGAAATAGACACTTTTCTACTGATCATATACTGTGTATCTGAGCAG	3111
Db	290	AAACCCATGCTGTGAAATAGACACTTTTCTACTGATCATATACTGTGTATCTGAGCAG	231
Qy	3112	TGATTCGCAACCCATCTGTAAGTCACACAGAAAGTCCAGGTTTAAATTCGTCGCGGATCT	3171
Db	230	TGATTCGCAACCCATCTGTAAGTCACACAGAAAGTCCAGGTTTAAATTCGTCGCGGATCT	171
Qy	3172	GTGCGATATCTGAAAAAATGAAACCGTAGTTTTTGTTTTTAAATACAGAAATCATGTTG	3231
Db	170	GTGCGATATCTGAAAAAATGAAACCGTAGTTTTTGTTTTTAAATACAGAAATCATGTTG	111
Qy	3232	TTTTCGCACTTATATATAAAGCATGGAAGAAATTTACTTAGAGCAATTGTAAACATT	3291
Db	110	TTTTCGCACTTATATATAAAGCATGGAAGAAATTTACTTAGAGCAATTGTAAACATT	51
Qy	3292	TTGGAAGTAAACCATTTGAGATTGAAATTAACGCGATAGTGGTTCCTT	3340
Db	50	TTGGAAGTAAACCATTTGAGATTGAAATTAACGCGATAGTGGTTCCTT	2

```

RESULT 13
LOCUS      AI926727/c
DEFINITION AI926727
            753 bp      mRNA      linear      EST 08-MAR-2000
            w050901.x1 NCI CGAP G4a4 Homo sapiens cDNA clone IMAGE:2458800 3'
            similar to SW:GALD_HUMAN C05682 CALDESMON ; contains element TAR1
            repetitive element , mRNA sequence.
AI926727
AI926727.1 GI:5662691
EST.
Homo sapiens (human)
Homo sapiens
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 753)
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgsbbs-remail.nih.gov
Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Greg Lennon, Ph.D.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/ILN at:
www.bio.lnln.gov/bbrp/image/image.html
Insert Length: 3763      Std Error: 0.00
Seq. primer: -400P from Gibco
High quality sequence stop: 396.
Location/Qualifiers
1..753

```

```

1. 753
local:ACM,aaal,aaal,aaal
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2458600"
/tissue_type="poorly differentiated adenocarcinoma with
sigmat ring cell features"
/lab_host="DH10B"
/clone_1ib="NCI CGAP G554"
/note="Organ: stomach; Vector: pCMV-Sport6; Site: 1: SalI;
Site: 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.69 Kb. Life Technologies catalog #:"

```

ORIGIN	11549-01"
Query Match	19.1%; Score 708.8; DB 9; Length 753;
Best Local Similarity	96.3%; Pred. No. 9.6e-113;
Matches 725; Conservative	0; Mismatches 28; Indels 0; Gaps 0;

ORIGIN	11549-011"
Query Match	19.1%; Score 708.8; DB 9; Length 753;
Best Local Similarity	96.3%; Pred. No. 9,66-113;
Matches 725; Conservative	0; Mismatches 28; Indels 0; Gaps 0;
QY	1082 GAAGAGAAAACAACTGAAAGCCACAGAGAAAACAGTGGTATATGTCAATTAAAAATGGGCG 1141
DB	753 GAGCAGCCAGACTTGTAGAGCCAGCAGCACAACATGTATATGTATACACAAATGGCCG 694
QY	1142 ATCAGTTCAGAAAGACCTTAAACAAGAGAGAGAGGAAACAAGGTTCCAGATGATTTCC 1201
DB	693 ATCAGTCAGAAAGAGCCTATTCAGAGAGAGAAAGAGGAAACAAGGTCAGATGATTTCC 634
QY	1202 CATCATGAAAAAGATGAAAGAGGAAACAAGAAAGAGCTGAGGAGAGAGGCAAGGTG 1261
DB	633 CATCATGAAAAAGATGAAAGAGGAAACAAGAAAGAGCTGAGGAGAGAGGCAAGGTG 574
QY	1262 GAACCAAGAGAAAGAAAGATTTAAAGCCAGCAAGCAAAAAGATTCGACATGACGA 1321
DB	573 GAACCAAGAGAAAGAAAGATTTAAAGCCAGCAAGCAAAAAGATTCGACATGACGA 514
QY	1322 GCAAGAAATTGAAGCAGAAAGAAAGAGAGCTCCCAAGAAAGAGAAAGAGAGAGCAGAA 1381
DB	513 GCAAGAAATTGAAGCAGAAAGAAAGAGAGCTCCCAAGAAAGAGAAAGAGAGAGCAGAA 454
QY	1382 GAGAGGCAAAAGATGAGAGGAGAAAGAAAGAGGCAAGAGAGAGAGAGAGATAG 1441
DB	453 GAGAGGCAAAAGATGAGAGGAGAAAGAAAGAGGCAAGAGAGAGAGAGAGATAG 394
QY	1442 GAGGAGAGAAAGAGGCAAGAGAGAGCAGAGATTAAGAGAGAAAGAGAAAGGCA 1501
DB	393 GAGGAGAGAGAAAGAGGCAAGAGAGAGCAGAGATTAAGAGAGAAAGAGAAAGGCA 334
QY	1502 GCAGAGAGAGAGCAGAGAGATTTAAAGAGAGAAACAAGAAAGGAGCAGAGAGAGCAAG 1561
DB	333 GCAGAGAGAGAGCAGAGAGATTTAAAGAGAGAAACAAGAAAGGAGCAGAGAGAGCAAG 274
QY	1562 GCCAGGCGCAGAGAGAGAAAGAAAGGCTTAAGTAAGACAGCAAAACGTAAACAAGACTA 1621
DB	273 GCCAGGCGCAGAGAGAGAAAGAAAGGCTTAAGTAAGTAAGACAGCAAAACGTAAACAAGACTA 214
QY	1622 GAAGAGAAAAAAACATGCCATGCAAGAGACAAAGATTAAGAGGAGAAAGGTAAGAACAGAA 1681
DB	213 GAAGAGAAAAAAACATGCCATGCAAGAGACCAAGATTAAGAGGAGAAAGGTAAGAACAGAA 154
QY	1682 ATAGAAAGGAAATGGGTAAATGAAAGAAACCAAGAGATTAAGCTTCAGACAGCTGTC 1741
DB	153 ATAGAAAGGAAATGGGTAAATGAAAGAAACCAAGAGATTAAGCTTCAGACAGCTGTC 94
QY	1742 CTAAGAAAGAGAGAGAAAGAGAGGAACTTAAAGTGCAGAGCTTAAAGAGAAAGGCTCAA 1801
DB	93 CTAAGAAAGAGAGAGAGAAAGAGAGGAACTTAAAGTGCAGAGCTTAAAGAGAAAGGCTCAA 34
QY	1802 GAAGACAAGCTTACCTTCAAAAAAGAGAGATC 1834
DB	33 GAAGACAAGCTTACCTTCAAAAAAGAGAGATC 1
RESULT 14	
AUI40904	
LOCUS	AUI40904 PLACE4 Homo sapiens cDNA clone PLACE4000466 5', mRNA
DEFINITION	sequence.
ACCESSION	AUI40904
VERSION	AUI40904.1 GI:11002425
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 799)

AUTHORS
Ota, T., Suzuki, Y., Saito, K., Iehli, S., Yamamoto, J., Sugiyama, T., Nishikawa, T., Nakamura, Y., Sugano, S., Masuno, Y., and Isegai, T.
HRI human cDNA project (Ota, T., Suzuki, Y., Saito, K., Iehli, S., Yamamoto, J., Sugiyama, T., Nishikawa, T., Nakamura, Y., Sugano, S., Masuno, Y., Isegai, T.)
Unpublished (2000)

JOURNAL
Contact: Takao Isegai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project, 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.

FEATURES
Location/Qualifiers
1..799
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE400466"
/tissue_type="Placenta"
/clone_lib="PLACE4"
/note="Vector: pME18SFL3"

ORIGIN

Query Match 19.1%; Score 708.8; DB 9; Length 799;
Best Local Similarity 99.4%; Pred. No. 9.5e-113;
Matches 721; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

490 GGAGGAGATGCGACTGGAAGCAAGAAAGATGCGCTACGAGAGATGACATGATGAGA 549
20 GGAGAGATGCGACTGGAAGCAAGAAAGATGCGCTACGAGAGATGACATGATGAGA 79.
550 GGAGGAG 609
80 GGAGGAG 139
610 GGAGAGATGCGACTGGAAGCAAGAAAGATGCGCTACGAGAGATGACATGATGAGA 669
140 GGAGAGATGCGACTGGAAGCAAGAAAGATGCGCTACGAGAGATGACATGATGAGA 199
670 TGAGGAG 729
200 TGAGGAG 259
730 ATTCTGAG 789
260 ATTCTGAG 319
790 GGAGGAG 849
320 GGAGGAG 379
850 AAGATGCAAAATGACACACAGCAAAATGAACTACCGAGAGAGAGAGAGAGAGAGAGAG 909
380 AAGATGCAAAATGACACACAGCAAAATGAACTACCGAGAGAGAGAGAGAGAGAGAGAG 439
910 TCGCCAAAGAAAGTACGAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 969
440 TCGCCAAAGAAAGTACGAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 499
970 TGAATGAG 1029
500 TGAATGAG 559
1030 GGAG 1089
560 GGAG 619
1090 AACCACTGAAG 1149

RESULT 15
BX642011
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

620 AACCACTGAAG 679
1150 AAG 1208
680 AAG 739
1209 AAAAG 1213
740 AAAAG 744

735 bp mRNA linear EST 04-SEP-2003
DKFZp686H6201_r1 686 (synonym: hicc3) Homo sapiens cDNA clone
DKFZp686H6201 5', mRNA sequence.
BX642011
BX642011 GI:34476330
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 735)
Poustka, A., Albert, R., Moosmayer, P., Schupp, I., Wellenreuther, R.,
Mewes, R.W., Well, B., Amd, C., Osanger, A., Fobo, G., Han, W. and
Wiemann, S.
EST (Poustka, A., Albert, R., Moosmayer, P., Schupp, I.,
Wellenreuther, R., et al.)
Unpublished (2003)
Contract: MIPS

FEATURES
source
1..735
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686H6201"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="686 (synonym: hicc3)"
/note="Vector: pTIP1EX2; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"

ORIGIN

Query Match 18.9%; Score 701.2; DB 13; Length 735;
Best Local Similarity 99.3%; Pred. No. 2e-111;
Matches 714; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

201 AGTGTATTCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTA 260
2 AGTGTATTCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTA 61
261 AAGATCAGTCCCTTCCTTCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTA 320
62 AAGATCAGTCCCTTCCTTCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTA 121
321 CATTTGAACATTTCAAGATCATCAATCAATTTCAAGAGAGAGAGAGAGAGAGAGAGAG 380
122 CATTTGAACATTTCAAGATCATCAATCAATTTCAAGAGAGAGAGAGAGAGAGAGAGAG 181

Fri Mar 12 10:35:45 2004

us-10-084-817-30.rst

Page 13

Oy	381	GCTCAGACATCTGATTCGACCTGTCCAGACATCATCTGATCTCCTGAACTGAAAT	440
Db	182	GCTCAGACATCTGATTCGACCTGTCCAGACATCATCTGATCTCCTGAACTGAAAT	241
Oy	441	ACACATGGATGATTTTGGAGCTTCGAGAGAACTTGGAAAGGCAGAAAGAGAGAAATGC	500
Db	242	ACACATGGATGATTTTGGAGCTTCGAGAGAACTTGGAAAGGCAGAAAGAGAGAAATGC	301
Oy	501	GACTCGAAGCAGAAAGAAATCGCCTTACAGAGGATACGATGATGAAAGAGAGAGCAAGCC	560
Db	302	GACTCGAAGCAGAAAGAAATCGCCTTACAGAGGATACGATGATGAAAGAGAGAGCAAGCC	361
Oy	561	GGAAACGGCGCCGCGAGCCCGACAGAAACGCGCTGCGCGAGAAAGCAGAGAGAAAGAAATCCT	620
Db	362	GGAAACGGCGCCGCGAGCCCGACAGAAACGCGCTGCGCGAGAAAGCAGAGAGAAAGAAATCCT	421
Oy	621	TGGGACAGGTGACCGACCGACAGGTGGAGGTGAATGCCAGAGAAAGATGTGTCCCTGACAGAGAGG	680
Db	422	TGGGACAGGTGACCGACCGACAGGTGGAGGTGAATGCCAGAGAAAGATGTGTCCCTGACAGAGAGG	481
Oy	681	CCAGACAAACCAACCAAAACACTCAAGTGGAGAGGGATGATGAGCCGCATTCCTTGAGC	740
Db	482	CCAAACAAACCAACCAAAACACTCAAGTGGAGAGGGATGATGAGCCGCATTCCTTGAGC	541
Oy	741	GCCTGGCTGGGCGTGAGAGAAAGCGCCAAAAAGCGCTTCAGAGAGCGCTTCGAGACGCGACA	800
Db	542	GCCTGGCTGGGCGTGAGAGAAAGCGCCAAAAAGCGCTTCAGAGAGCGCTTCGAGACGCGACA	601
Oy	801	AGGAGTTGGAACCCCAACATTAACGATGCGAAGTGTGTGCTCCCAAGAGAGAAATGC	860
Db	602	AGGAGTTGGAACCCCAACATTAACGATGCGAAGTGTGTGCTCCCAAGAGAGAAATGC	661
Oy	861	ATGACACAGCAGAAATGAAATCTACCGGAGAGAGAAAGAAAGATGAAAGTCCCGCAAGA	919
Db	662	ATGACACAGCAGAAATGAAATCTACCGGAGAGAGAAAGAAAGATGAAAGTCCCGCAAGA	719

Search completed: March 11, 2004, 14:23:47
Job time : 6135.29 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 10, 2004, 08:38:17 ; Search time 5233.4 Seconds
(without alignments)
17011.249 Million cell updates/sec

Title: US-10-084-817-32
Perfect score: 2054
Sequence: 1 ccggccctgcctgctccgc.....cattgaaaaaaaaaaaaa 2054

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_btg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2037.2	99.2	2043	9 BC031082	BC031082 Homo sapi
2	2034.4	99.0	2036	6 D87258	D87258 Homo sapien
3	2034.4	99.0	2075	6 E12965	E12965 cDNA encodi
4	2034	99.0	2036	6 AR095627	AR095627 Sequence
5	2025	98.6	2029	9 HSSERIPRT	Y07921 H.sapiens m
6	1766.6	86.0	1900	9 AK092476	AK092476 Homo sapi
7	1701.4	82.8	1894	6 AX747270	AX747270 Sequence
8	1701.4	82.8	1894	9 AK091944	AK091944 Homo sapi
9	1644	80.0	1705	6 AX015404	AX015404 Sequence
10	1644	80.0	1705	6 BD194550	BD194550 Human nuc
11	1455.2	70.8	1534	6 AX202090	AX202090 Sequence
12	1397	68.0	1423	9 AF070555	AF070555 Homo sapi
13	1330.2	64.8	2023	10 BC013516	BC013516 Mus muscu
14	1321	64.3	2005	10 AF179369	AF179369 Mus muscu
15	1321	64.3	2007	10 AK128916	AK128916 Mus muscu
16	1302.4	63.4	2010	10 AF172994	AF172994 Mus muscu
17	1301.8	63.4	2012	10 AF179370	AF179370 Rattus no
18	1162.6	56.6	1169	9 AF097709	AF097709 Homo sapi
19	953	46.4	1113	4 AF097707	AF097707 Bos tauru
20	717	34.9	79597	9 AF157623	AF157623 Homo sapi
21	717	34.9	111002	9 EX842242	EX842242 Homo sapi
22	717	34.9	183623	2 AL844554	AL844554 Mus muscu
23	717	34.9	191741	9 AL603764	AL603764 Human DNA
24	533.4	26.0	604	6 BD222460	BD222460 Drug-rela
25	493.4	24.0	495	6 AR273154	AR273154 Sequence
26	493.4	24.0	495	6 AR276735	AR276735 Sequence
27	493.4	24.0	495	6 AR407010	AR407010 Sequence
28	493.4	24.0	495	6 AX368187	AX368187 Sequence
29	479.4	23.3	539	6 I74320	I74320 Sequence 18
30	451.6	22.0	1797	6 AX338509	AX338509 Sequence
31	451.6	22.0	2543	9 AY280665	AY280665 Homo sapi
32	451.6	22.0	2554	9 AY040094	AY040094 Homo sapi
33	451.6	22.0	2585	9 BC034390	BC034390 Homo sapi
34	451.6	22.0	2586	9 BC035717	BC035717 Homo sapi
35	427.4	20.8	2446	10 AY156509	AY156509 Mus muscu
36	426.2	20.7	519	10 AF097706	AF097706 Cavia por
37	419	20.4	441	6 AR422696	AR422696 Sequence
38	419	20.4	441	6 BD118249	BD118249 Esf and e
39	414.6	20.2	1521	9 BC057765	BC057765 Homo sapi
40	414.6	20.2	1544	6 BD127576	BD127576 Primer fo
41	414.6	20.2	1544	9 AK075205	AK075205 Homo sapi
42	414.6	20.2	1553	6 BD063243	BD063243 Secreted
43	406.6	19.8	2040	6 BD260123	BD260123 Matrix-re
44	397.8	19.4	489	4 AF097708	AF097708 Oryctolag
45	391.4	19.1	1953	9 AY280666	AY280666 Homo sapi

ALIGNMENTS

RESULT 1
EC031082
LOCUS
DEFINITION Homo sapiens protease, serine, 11 (IGF binding), mRNA (cDNA clone MGC:33808 IMAGE:5289018), complete cds.
ACCESSION BC031082
VERSION BC031082.1 GI:21410568
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2043)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,

841	Db		CGCTCCTCAGAGCTGCGCGCGGAGAGTTCGTGGTCCCATCGGAAGCCCGTTTTCCCTT	900
913	Qy	CAAAA	CACAGTCACCAACCCGGATTCGTGAGCACCAACCAGCGAGCGCGGCAAGAGCTGGGG	972
901	Db	CAAAA	CACAGTCACCAACCCGGATTCGTGAGCACCAACCAGCGAGCGCGGCAAGAGCTGGGG	960
973	Qy	CTCCG	CAACTCAGACATGACATCATCCAGACCGACGCCATCATCAACTATGGAACTCG	1032
961	Db	CTCCG	CAACTCAGACATGACATCATCCAGACCGACGCCATCATCAACTATGGAACTCG	1020
1033	Qy	GGAGG	CCGCTTAGTAAACCTGGACGGTGAAGTGATTTGGAATTAACACTTTGAAAGTGACA	1092
1021	Db	GGAGG	CCGCTTAGTAAACCTGGACGGTGAAGTGATTTGGAATTAACACTTTGAAAGTGACA	1080
1093	Qy	GCTGG	AATCTCTTTTGCATCCCATCTGATAAGATTAAGAGTTCTCTCAGGAGTCCCAT	1152
1081	Db	GCTGG	AATCTCTTTTGCATCCCATCTGATAAGATTAAGAGTTCTCTCAGGAGTCCCAT	1140
1153	Qy	GACCG	CAGGCCAAAGGAAAAGCCATCACCAAGAAAGTATATTTGGTATCCGAATGATG	1212
1141	Db	GACCG	CAGGCCAAAGGAAAAGCCATCACCAAGAAAGTATATTTGGTATCCGAATGATG	1200
1213	Qy	TCAC	TCCAGCCAGCAAGCCCAAGAGCTGAAGGACCGCACCGGACATCCACAGACGTG	1272
1201	Db	TCAC	TCCAGCCAGCAAGCCCAAGAGCTGAAGGACCGCACCGGACATCCACAGACGTG	1260
1273	Qy	ATCTC	AGGCGGTATATAATTGAAGTAATTCCTGATACCCACAGACAGAGTGGTGGTCTC	1332
1261	Db	ATCTC	AGGCGGTATATAATTGAAGTAATTCCTGATACCCACAGACAGAGTGGTGGTCTC	1320
1333	Qy	AAGG	AAAAAGCAGCTCATATACGATCAATGGACGTCCGTGGTCTCCGCGCAATGATGTC	1392
1321	Db	AAGG	AAAAAGCAGCTCATATACGATCAATGGACGTCCGTGGTCTCCGCGCAATGATGTC	1380
1393	Qy	AGCG	AGTCATTAAAGGGAAAGCAACCTGAAATCATGGTGGTCCGACGGGTAAATGAAGAT	1452
1381	Db	AGCG	AGTCATTAAAGGGAAAGCAACCTGAAATCATGGTGGTCCGACGGGTAAATGAAGAT	1440
1453	Qy	ATCAT	GATCACAGTGATTCGGAAGAAATTTGACCCATAGGCAGAGGCATGAGCTGACATT	1512
1441	Db	ATCAT	GATCACAGTGATTCGGAAGAAATTTGACCCATAGGCAGAGGCATGAGCTGACATT	1500
1513	Qy	CATGT	TTCCCTCAAGACTCTCCGCTGGATGACGGATGAGGACTCTGGGCTGTGGAATA	1572
1501	Db	CATGT	TTCCCTCAAGACTCTCCGCTGGATGACGGATGAGGACTCTGGGCTGTGGAATA	1560
1573	Qy	GGAC	ACTCAAGACTTTTGACTGCAATTTCTTTGTTTCAGTGGAGACTCCCTGCGCCAAACAG	1632
1561	Db	GGAC	ACTCAAGACTTTTGACTGCAATTTCTTTGTTTCAGTGGAGACTCCCTGCGCCAAACAG	1620
1633	Qy	AATCC	TTCTTGTATGATGATGACAGTGTGCTTTTCTTGCCAGCTTTGGGCCATTCT	1692
1621	Db	AATCC	TTCTTGTATGATGATGACAGTGTGCTTTTCTTGCCAGCTTTGGGCCATTCT	1680
1693	Qy	CTCTG	CCCTTCTGTATCCTATGATGACAGTGTGCTTTTCTTGCCAGCTTTGGGCCATTCT	1752
1681	Db	CTCTG	CCCTTCTGTATCCTATGATGACAGTGTGCTTTTCTTGCCAGCTTTGGGCCATTCT	1740
1753	Qy	TGCTT	AGACAGTCAGCATTTGTCTCTCTTTAACTGAGTCATCATCTTAGTCCCACTAA	1812
1741	Db	TGCTT	AGACAGTCAGCATTTGTCTCTCTTTAACTGAGTCATCATCTTAGTCCCACTAA	1800
1813	Qy	TGCAG	TGATACATGCGGTAGATAGAAAGCCCAACGGAGCCAGGATGGGACTTGCTCG	1872
1801	Db	TGCAG	TGATACATGCGGTAGATAGAAAGCCCAACGGAGCCAGGATGGGACTTGCTCG	1860
1873	Qy	TGTTT	GTGCTTTTCTCCAGTCAGCACCCAAAGTCAATGCAACAGACACCCGGTGGGT	1932
1861	Db	TGTTT	GTGCTTTTCTCCAGTCAGCACCCAAAGTCAATGCAACAGACACCCGGTGGGT	1920
1933	Qy	GAGCG	CTGGCTTCTCAACCGGCCGAAGTGTGCTCTTTTAGGAATCTCTTTTGGAAATTGGGA	1992

Db	1921	GAGCGTGGCTTCTCAACAGCGCGAAGTTGCCTCTCTTTTAGGAATCTCTTTTGAATGGGA	1980
Qy	1993	GCACGATGACTCTGAGTTTGAGCTATTAAAGTACTTCTTACACATTGAAAAA	2052
Db	1981	GCACGATGACTCTGAGTTTGAGCTATTAAAGTACTTCTTACACATTGAAAAA	2040
Qy	2053	AA	2054
Db	2041	AA	2042
RESULT 2			
LOCUS	D87258		
DEFINITION	Homo sapiens mRNA for serin protease with IGF-binding motif, complete cds.	2036 bp	linear
ACCESSION	D87258		
VERSION	D87258.1	GI:1513058	
KEYWORDS	serin protease with IGF-binding motif.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Ono, I., Hashimoto, J., Takaoka, K., Takahiro, O., Okubo, K. and Matsubara, K.		
TITLE	The cloning of a cDNA for novel serine protease with IGF-binding motif expressed in human osteoblast		
JOURNAL	Unpublished		
AUTHORS	Ono, I. to 2036		
TITLE	Direct Submission		
JOURNAL	Submitted (20-AUG-1996) Ikko Onno, Institute for Molecular and Cellular Biology, Osaka University, Molecular Genetics; 1-3 Yamada-oka, Suita, Osaka 565, Japan		
FEATURES	(E-mail:kousaku@imcb.osaka-u.ac.jp, Tel:81-6-879-7992, Fax:81-6-877-1922)		
source	Location/Qualifiers		
	1. 2036		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="GS2422"		
	/cell_type="osteoblast"		
	/tissue_type="cancellous bone"		
	/clone_lib="lambda ZAP"		
	/tissue_lib="3 end-directed library"		
	49..1451		
	/note="GS2422 protein"		
	/codon_start=1		
	/product="serin protease with IGF-binding motif"		
	/protein_id="BA13322.1"		
	/db_xref="GI:1513059"		
CDS			
	/translation="MQIPRAALPLLLLLAASPAAQLSRAGRSAPLAAGCDPCRCBPA RCPPQCEGGRADACCCCEVCGAGACGLQEGPCGRLQCVFPGVPASATVR RAQAGLCVCSSEPCVCGSDANYANLCQRAARRSERLHRPPVIVLQRCAGCGQGE DPNSLRHKVFTADVVEKIPAPAVVHIEFLPKLPFSKREVPVASGSGFVISEDGLIVTN AHVNTNHRVKNFLKNGATYEAIKQDVDEKADIALIKDHQKGLPVLLGRSSELRLP EFVVAIGSPFLQNTVTTGIVSTQRCGKELGLENMDYIQTDIAIINYNGSGPLVN LDGEVAGINTSLQVATIGISPAISDKIKKFLTEHSDROAKGAIKKTKYIGIRMSLT SKAKELKDRHDPDPVISAQIIEVTPDTPAEGAGLKNVDVVISNGQSWASANDVSD VTKPESTNMVPEGNFTWTVIPEIDP"		

ORIGIN

```
Query Match          99.0%; Score 2034.4; DB 9; Length 2036;
Best Local Similarity 100.0%; Pred.No. 4.8e-231;
Matches 2035; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

Qy 1 CGGGCCTCGCCGTGTCGCCGCCACCGCGCGCGGCAGATCGCCATGCAgATCCG 60
|||||
|||||
Db 1 CGGGCCTCGCCGTGTCGCCGCCACCGCGCGCGGCAGATCGCCATGCAgATCCG 60
|||||
61 CGGGCCTGTTCTCCGCTGTGTGCTGTACTGGCGGGCGCCGCGTCCGGCGCAGCTG 120

[illegible]


```
Db 485 CTGCGAGCGGAGCGCTCGGCCAAGGCGAGAGATCCCAACAGTTTGGGCCATAAATA 544
QY 555 TAACTTTATCGCGAGCGTGGTGGAGAGATCGCCCTCGCGTGGTTCATATCGAATTGTT 614
Db 545 TAACTTTATCGCGAGCGTGGTGGAGAGATCGCCCTCGCGTGGTTCATATCGAATTGTT 604
QY 615 TCGCAAGCTTCCGTTTTCTAAACGAGAGAGTCCCGTGGCTAGTGGGTCTGGGTTTATTGT 674
Db 605 TCGCAAGCTTCCGTTTTCTAAACGAGAGAGTCCCGTGGCTAGTGGGTCTGGGTTTATTGT 664
QY 675 GTCGGAAGATGGAGCTGATCGTGACAAATGCCACGCTGGTGACCAACAAGCACC GGGTCAA 734
Db 665 GTCGGAAGATGGAGCTGATCGTGACAAATGCCACGCTGGTGACCAACAAGCACC GGGTCAA 724
QY 735 AGTTGAGCTGAAGAACGGTGCCTATTACGAGGCCAAATCAAGAGATGGGATGAGAAAGC 794
Db 725 AGTTGAGCTGAAGAACGGTGCCTATTACGAGGCCAAATCAAGAGATGGGATGAGAAAGC 784
QY 795 AGACATCGCACTCATCAAAATTTGACCAACGAGGCAAGTGCCTGCTCTGCTGGCG 854
Db 785 AGACATCGCACTCATCAAAATTTGACCAACGAGGCAAGTGCCTGCTCTGCTGGCG 844
QY 855 CTGCTAGAGCTGGCGCGGGAGAGTTCGTGGTGCCTATCGGAAGCCGTTTTCCCTTCA 914
Db 845 CTGCTAGAGCTGGCGCGGGAGAGTTCGTGGTGCCTATCGGAAGCCGTTTTCCCTTCA 904
QY 915 AAACACAGTCAACACCGGGATCGTGAGCACCACCGAGGCGGCAAGAGCTGGGGCT 974
Db 905 AAACACAGTCAACACCGGGATCGTGAGCACCACCGAGGCGGCAAGAGCTGGGGCT 964
QY 975 CCGCACTGACATGACATCACTCCAGCAGCGCCATCATCACTATGGAATCGCG 1034
Db 1024 CCGCACTGACATGACATCACTCCAGCAGCGCCATCATCACTATGGAATCGCG 1024
QY 1035 AGCCCGCTTAACTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1094
Db 1025 AGCCCGCTTAACTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1084
QY 1095 TGGATCTCTTTGCAATCCCATCTGATGAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1154
Db 1085 TGGATCTCTTTGCAATCCCATCTGATGAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1144
QY 1155 CCGACAGCCAAAGGAAAGCCATCACAAGAAAGTATATGTTATCCGATGATGTC 1214
Db 1145 CCGACAGCCAAAGGAAAGCCATCACAAGAAAGTATATGTTATCCGATGATGTC 1204
QY 1215 ACTCAGCTCCAGCAAGGCAAGAGCTCAAGGACCGGACCGGACCTCCAGACGTGAT 1274
Db 1205 ACTCAGCTCCAGCAAGGCAAGAGCTCAAGGACCGGACCGGACCTCCAGACGTGAT 1264
QY 1275 CTCAGGAGCGTATATATGAGTAATTCCTGATACCCAGCAGAGCTGGTGGTCAA 1334
Db 1265 CTCAGGAGCGTATATATGAGTAATTCCTGATACCCAGCAGAGCTGGTGGTCAA 1324
QY 1335 GGAAACGACGTCATAATCAGCATCAATGGACAGTCCGTTGCTCCGCCAATGATCTAG 1394
Db 1325 GGAAACGACGTCATAATCAGCATCAATGGACAGTCCGTTGCTCCGCCAATGATCTAG 1384
QY 1395 CGAGCTATTAAGGGAAGCAGCCTGACATGTTGTTCCGACGGGTGAATGAAGATAT 1454
Db 1385 CGAGCTATTAAGGGAAGCAGCCTGACATGTTGTTCCGACGGGTGAATGAAGATAT 1444
QY 1455 CATGATCAGATGATTCGCAAGAAATGACCCATAGGAGGAGGATGAGCTGATTCAT 1514
Db 1445 CATGATCAGATGATTCGCAAGAAATGACCCATAGGAGGAGGATGAGCTGATTCAT 1504
QY 1515 TGTTCCTCAAGACTCTCCGCTGATGACGATGAGGACTCTGGGCTGCTGGAATAGG 1574
Db 1505 TGTTCCTCAAGACTCTCCGCTGATGACGATGAGGACTCTGGGCTGCTGGAATAGG 1564
QY 1575 ACATCAAGACTTTGATCTGCCATTTTGTGTTTCACTGAGGACTCCCTGGCCCAAGAA 1634
Db 1565 ACATCAAGACTTTGATCTGCCATTTTGTGTTTCACTGAGGACTCCCTGGCCCAAGAA 1624
```

```
QY 1635 TCCTTCTTGATAGTTTCAGGCAAAAACAAATGTAATGTTGCAGATCCGCGAGCAGAAGCT 1694
Db 1625 TCCTTCTTGATAGTTTCAGGCAAAAACAAATGTAATGTTGCAGATCCGCGAGCAGAAGCT 1684
QY 1695 CTGCGCTTCTGATCCTATGATGAGTGTGCTTTTCTTGCAGCTTGGGCAATCTTG 1754
Db 1685 CTGCGCTTCTGATCCTATGATGAGTGTGCTTTTCTTGCAGCTTGGGCAATCTTG 1744
QY 1755 CTTAGACAGTCAGCATTTGTCCTCTCTTAACTGAGTCATCATCTTAGTCCAACATAATG 1814
Db 1745 CTTAGACAGTCAGCATTTGTCCTCTCTTAACTGAGTCATCATCTTAGTCCAACATAATG 1804
QY 1815 CAGTCGATCAATCGCTAGTAGAGAGAGCCCGGAGCCAGGATGGAGCTGGTGGTG 1874
Db 1805 CAGTCGATCAATCGCTAGTAGAGAGAGCCCGGAGCCAGGATGGAGCTGGTGGTG 1864
QY 1875 TTTGTGCTTTTCTCCAAGTCAGCACCAGGTCAATGCACAGAGACCCCGGCTGGTGA 1934
Db 1865 TTTGTGCTTTTCTCCAAGTCAGCACCAGGTCAATGCACAGAGACCCCGGCTGGTGA 1924
QY 1935 GCGCTGCTTCTCAACGCGCGAAGTTGCTCTTTTAGGAATCTCTTTGGAATGGGAGC 1994
Db 1925 GCGCTGCTTCTCAACGCGCGAAGTTGCTCTTTTAGGAATCTCTTTGGAATGGGAGC 1984
QY 1995 ACGATGACTGAGTTTGAAGTATTAAGTACTTCTTACACATG 2039
Db 1985 ACGATGACTGAGTTTGAAGTATTAAGTACTTCTTACACATG 2029
```

```
RESULT 6
AK092476 1900 bp mRNA linear PRI 15-JUL-2002
LOCUS Homo sapiens cDNA FLJ35157 fis, clone PLACE011156, highly similar
DEFINITION AK092476
ACCESSION AK092476
VERSION AK092476.1 GI:21751082
KEYWORDS oligo capping; fis (full insert sequence)
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
```

REFERENCE 1
AUTHORS Kawakami, B., Sugiyama, A., Takemoto, M., Sugiyama, T., Irie, R., Oseuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Suzuki, Y., Sugano, S., Nagahara, K., Masuho, Y., Nagai, K. and Isogai, T.

TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1900)
AUTHORS Isogai, T. and Yamamoto, J.
JOURNAL Direct Submission

TITLE Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES
source
1. 1900
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="PLACE011156"
/tissue_type="placenta"

```
/clone_lib="PLACE6"  
/note="cloning vector: pME18SFL3"
```

ORIGIN

Query Match	86.0%;	Score 1766.6;	DB 9;	Length 1900;
Best Local Similarity	99.2%;	Pred. No. 1.8e-251;		
Matches 1775;	Conservative 0;	Mismatches 14;	Indels 0;	Gaps 0;
QY	251	GC	GC	GC
DB	112	GC	GC	GC
QY	311	AG	TC	GC
DB	172	AG	TC	GC
QY	371	GC	TC	GC
DB	232	GC	TC	GC
QY	431	AC	TC	GC
DB	292	AC	TC	GC
QY	491	TC	GC	TC
DB	352	TC	GC	TC
QY	551	AA	TA	TA
DB	412	AA	TA	TA
QY	611	TG	TT	TC
DB	472	TG	TT	TC
QY	671	TT	GC	GC
DB	532	TT	GC	GC
QY	731	TC	AA	GT
DB	592	TC	AA	GT
QY	791	AA	GC	AC
DB	652	AA	GC	AC
QY	851	GC	GC	TC
DB	712	GC	GC	TC
QY	911	TT	CA	AA
DB	772	TT	CA	AA
QY	971	GC	TC	GC
DB	832	GC	TC	GC
QY	1031	CG	GC	GC
DB	892	CG	GC	GC
QY	1091	CAG	CT	GC
DB	952	CAG	CT	GC
QY	1151	AT	GC	GC
DB	1012	AT	GC	GC
QY	1211	TG	TC	GC

RESULT 7

AX747270	AX747270	1894 bp	linear	PAT 20-JUN-2003
LOCUS	Sequence 795 from Patent EP1308459.			
DEFINITION	AX747270			
ACCESSION	AX747270.1	GI:32131658		
VERSION				
KEYWORDS				
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1			
AUTHORS	Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ihii, S., Yamamoto, J. I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R., Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahari, K. and Masuho, Y.			
TITLE	Full-length cDNA sequences			


```
QY 1767 GCATTGTCCTCCTTAACTGAGTCATCATCTTAGTCCCACTAATGCAGTCGATACAA 1826
Db 1379 GCAATTGTCCTCCTTAACTGAGTCATCATCTTAGTCCCACTAATGCAGTCGATACAA 1438
QY 1827 TGGTATAGTAAAGAACCCACCGGAGCCAGGATGGGACTGGTCTGTTGCTTTTC 1886
Db 1439 TGGTATAGTAAAGAACCCACCGGAGCCAGGATGGGACTGGTCTGTTGCTTTTC 1498
QY 1887 TCCAAAGTCACACCAAGGTCNAATGCACAGACCCCGGTGGGTGAGCGTGGCTTCT 1946
Db 1499 TCCAAAGTCACACCAAGGTCNAATGCACAGACCCCGGTGGGTGAGCGTGGCTTCT 1558
QY 1947 CAAACCGCCGAAGTGCCTCTTTTAGGAATCTCTTTGGAATGGGAGCAGATGACTCTG 2006
Db 1559 CAAACCGCCGAAGTGCCTCTTTTAGGAATCTCTTTGGAATGGGAGCAGATGACTCTG 1618
QY 2007 AGTTTGAGCTATTAAAGTACTCTTACACATTGAAATGAAAAA 2054
Db 1619 AGTTTGAGCTATTAAAGTACTCTTACACATTGAAATGAAAAA 1666

RESULT 10
BD194550
LOCUS Human nucleic acid sequence originating in normal ovarian tissue.
DEFINITION BD194550
ACCESSION BD194550.1 GI:33004292
KEYWORDS JP 2002510486-A/61.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1705)
Speft,I., Hintzman,B., Shemitt,A., Pirarski,C., Duhl,E. and
Rosenthal,A.
Human nucleic acid sequence originating in normal ovarian tissue
Patent: JP 2002510486-A 61 09-APR-2002;
METAGEN GESISCHAFT FUER GENOME FORSCHUNG MBH
OS Homo sapiens (human)
PN JP 2002510486-A/61
PD 09-APR-2002
PF 31-MAR-1999 JP 2000542440
PR 03-APR-1998 DE 198 16 395.9
PI THOMAS SPEFT, BERND HINTZMAN, ARMIN SHCMITT, CHRISTIAN PIRARSKI,
PI EDGAR DUHL,
PI ANDRE ROSENTHAL
PC C12N15/09,A61K38/00,A61K48/00,A61P35/00,C07K14/47,C07K16/30,
PC C12N1/15,C12N1/21,C12N5/10,C12Q1/68,G01N33/574,C12N15/00, PC
PC C12N1/19,C12N1/21,C12N5/10,C12Q1/68,G01N33/574,C12N15/00, PC
A61K37/02,
PC C12N5/00
CC Human nucleic acid sequence originating in normal ovarian CC
tissue
FH Key Location/Qualifiers
FT source 1..1705
FT /organism='Homo sapiens (human)'.
FEATURES
source
1..1705
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'

ORIGIN
Query Match 80.0%; Score 1644; DB 6; Length 1705;
Best Local Similarity 99.9%; Pred. No. 2.3e-233;
Matches 1666; Conservative 0; Mismatches 0; Indels 2; Gaps 2;
QY 387 CAGCAGGAGCGCGTGTGCGGAGCGGAGCCCAACACCTACGCCAACCTGTGCCAGCTGG 446
Db 1 CAGCAGGAGCGCGTGTGCGGAGCGGAGCCCAACACCTACGCCAACCTGTGCCAGCTGG 59
QY 447 CGCGCCAGCGCGCTCCGAGAGCTGCACCGCGCGCGGTCTGCTCTGCTGAGCGCGG 506
Db 1139 AGACTCTCCCTGGATGACGGATGAGGACTCTGGGCTGCTGGATAGGACACTCAAGACT 1198
```

```
Db 60 CGCGCCAGCGCGCTCCGAGAGGCTGCACCGCGCGCGGTCTCATGTCCTGCGAGCGGG 119
QY 507 AGCCTGCGGCCCAAGGCGAGGAGATCCCAAAGTTTGGCCCATTAATATATATATATATAT 566
Db 120 AGCCTGCGGCCCAAGGCGAGGAGATCCCAAAGTTTGGCCCATTAATATATATATATATAT 179
QY 567 GGAGCTGCTGAGAGAGATCGCCCTCGCGTGGTTCATATATCGAATTTGTTTCGCAAGCTTCC 626
Db 180 GGAGCTGCTGAGAGAGATCGCCCTCGCGTGGTTCATATATCGAATTTGTTTCGCAAGCTTCC 239
QY 627 GTTTTCTAAACGAGAGGTCGCGTGGTCTAGTGGGCTTGGGTTTATTTGTTGCGAAGATGG 686
Db 240 GTTTTCTAAACGAGAGGTCGCGTGGTCTAGTGGGCTTGGGTTTATTTGTTGCGAAGATGG 299
QY 687 ACTGATCGTGCACAAATGCCACGTCGTGACCAACAAAGCACCGGGTCAAAGTTGAGCTGAA 746
Db 300 ACTGATCGTGCACAAATGCCACGTCGTGACCAACAAAGCACCGGGTCAAAGTTGAGCTGAA 359
QY 747 GAACGCTGCCACTTACGAGGCCAAATCAAGATGTGGATGAGAAAGCAGACATCGCACT 806
Db 360 GAACGCTGCCACTTACGAGGCCAAATCAAGATGTGGATGAGAAAGCAGACATCGCACT 419
QY 807 CATCAAAATTCACCAACGAGGCAAGCTGCTCTGCTGCTTGGCGGCTCTCTCAGAGCT 866
Db 420 CATCAAAATTCACCAACGAGGCAAGCTGCTCTGCTGCTTGGCGGCTCTCTCAGAGCT 479
QY 867 GCGGCCGCGGAGAGTTCGTGTCGCCATCGGAAGCCGCTTTCCCTTCAAACAACAGTCAC 926
Db 480 GCGGCCGCGGAGAGTTCGTGTCGCCATCGGAAGCCGCTTTCCCTTCAAACAACAGTCAC 539
QY 927 CACCGGATCGTGAAGCACCCACGAGGCGGCAAGAGCTGGGGCTCCGCAACTCAGA 986
Db 540 CACCGGATCGTGAAGCACCCACGAGGCGGCAAGAGCTGGGGCTCCGCAACTCAGA 599
QY 987 CATGGAATCATCCAGACCGCCATCATCAACTATGGAATCGGGAGGCGCGTTAGT 1046
Db 600 CATGGAATCATCCAGACCGCCATCATCAACTATGGAATCGGGAGGCGCGTTAGT 659
QY 1047 AAACCTGGACGCTGAAGTGATTGGAATTAACACTTTGAAAGTGACAGCTGGAATCTCCTT 1106
Db 660 AAACCTGGACGCTGAAGTGATTGGAATTAACACTTTGAAAGTGACAGCTGGAATCTCCTT 719
QY 1107 TGCAATCCCATCTGATAAGATTAAAGTTCTCCAGGAGTCCCATGACCCGACAGGCCAA 1166
Db 720 TGCAATCCCATCTGATAAGATTAAAGTTCTCCAGGAGTCCCATGACCCGACAGGCCAA 779
QY 1167 AGGAAAGCCATCACCAAGAGAGTATATTGGTATCCGAATGATGTCCTCAGCTCCAG 1226
Db 780 AGGAAAGCCATCACCAAGAGAGTATATTGGTATCCGAATGATGTCCTCAGCTCCAG 839
QY 1227 CAAAGCCAAAGAGCTGAAGGACCGGACCTTCCAGAGCTGATCTCAGAGAGGTA 1286
Db 840 CAAAGCCAAAGAGCTGAAGGACCGGACCTTCCAGAGCTGATCTCAGAGAGGTA 899
QY 1287 TATAATTGAAGTAATTCCTGATACCCAGCAGAGCTGGTCTCAAGGAAACACGCT 1346
Db 900 TATAATTGAAGTAATTCCTGATACCCAGCAGAGCTGGTCTCAAGGAAACACGCT 959
QY 1347 CATTAATCAGCATCAATGGACAGTCCGTCCTCCGCAATGATGTGAGGACGCTCATTA 1406
Db 960 CATTAATCAGCATCAATGGACAGTCCGTCCTCCGCAATGATGTGAGGACGCTCATTA 1019
QY 1407 AAGGAAAGCACCCTGACATGGTGGTCCGAGGAGTATGAGATATCATGATCAGCT 1466
Db 1020 AAGGAAAGCACCCTGAAACATGGTGGTCCGCA-GGGTAATGAGATATCATGATCAGCT 1078
QY 1467 GATTCCCGAAGAAATTGACCCATAGGAGGAGCATGAGCTGGACTTTTCCCTCAA 1526
Db 1079 GATTCCCGAAGAAATTGACCCATAGGAGGAGCATGAGCTGGACTTTTCCCTCAA 1138
QY 1527 AGACTCTCCCTGGATGACGGATGAGGACTCTGGGCTGCTGGATAGGACACTCAAGACT 1586
Db 1139 AGACTCTCCCTGGATGACGGATGAGGACTCTGGGCTGCTGGATAGGACACTCAAGACT 1198
```

```
QY 1587 TTGACTGCCAATTTTGTGTTTCAAGTGGAGACTCCCTGGCCAAACAGAAATCCTTCTTGATA 1646
Db 1199 TTTGACTGCCAATTTTGTGTTTCAAGTGGAGACTCCCTGGCCAAACAGAAATCCTTCTTGATA 1258
QY 1647 GTTTGCAGGCAAAACAAATGTAATGTTGAGATCCCGCAGCAGCAAGCTCTGCCCTTCTGT 1706
Db 1259 GTTTGCAGGCAAAACAAATGTAATGTTGAGATCCCGCAGCAGCAAGCTCTGCCCTTCTGT 1318
QY 1707 ATCTATATGATGCAATGCTGCTTTTCTTCCAGCTTGGGCAATCTTGTAGACAGTCA 1766
Db 1319 ATCTATATGATGCAATGCTGCTTTTCTTCCAGCTTGGGCAATCTTGTGTAGACAGTCA 1378
QY 1767 GCATTGTCTCTCTCTTAACTGAGTCATCATCTTAGTCCAACTAATGCAAGTGCATACAA 1826
Db 1379 GCATTGTCTCTCTCTTAACTGAGTCATCATCTTAGTCCAACTAATGCAAGTGCATACAA 1438
QY 1827 TCGGTAGATAGAAAGCCCAACGGAGCAGATGGAGCTGTGCTGTTGTGCTTTTC 1886
Db 1439 TCGGTAGATAGAAAGCCCAACGGAGCAGATGGAGCTGTGCTGTTGTGCTTTTC 1498
QY 1887 TCCAAGTCAGACCCAAAGGTCATATGCAAGACACCCCGGTCGGTGAGCGCTGGCTTCT 1946
Db 1499 TCCAAGTCAGACCCAAAGGTCATATGCAAGACACCCCGGTCGGTGAGCGCTGGCTTCT 1558
QY 1947 CAACGGCCAGTGGCTCTTTTAGGAATCTCTTTGGAAATGGAGCAGATGACTCTG 2006
Db 1559 CAACGGCCAGTGGCTCTTTTAGGAATCTCTTTGGAAATGGAGCAGATGACTCTG 1618
QY 2007 AGTTTGAGCTATTAAAGTACTTCTTACATTTGAAATGAAAAA 2054
Db 1619 AGTTTGAGCTATTAAAGTACTTCTTACATTTGAAATGAAAAA 1666

RESULT 11
LOCUS AX202090 1534 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 20 from Patent WO0153531.
ACCESSION AX202090
VERSION AX202090.1 GI:15391876
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Phippard,D., Vasanthakumari,G., Dotson,S. and Ma,X.J.
TITLE Osteoarthritis tissue derived nucleic acids, polypeptides, vectors,
and cells
JOURNAL Patent: WO 0153531-A 20 26-JUL-2001;
Pharmacia Corporation (US)
FEATURES
source
1..1534
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 70.8%; Score 1455.2; DB 6; Length 1534;
Best Local Similarity 99.3%; Pred. No. 1.8e-205;
Matches 1524; Conservative 0; Mismatches 3; Indels 7; Gaps 6;
QY 525 GGAAGATCCCAACAGTTTGGCCATAAATAACTTTATCCGGACCGTGGTGGAGAGAT 584
Db 1 GGAAGATCCCAACAGTTTGGCCATAAATAACTTTATCCGGACCGTGGTGGAGAGAT 60
QY 585 CGCCCTGCGGTGTTTCATATCGAATTTGTTTCGCAAGCTTCCGTTTCTAAACGAGAGGT 644
Db 61 CGCCCTGCGGTGTTTCATATCGAATTTGTTTCGCAAGCTTCCGTTTCTAAACGAGAGGT 120
QY 645 GCGGTGGTACTGGCTCTGGTTTATTTGTCGGAGATGGACTGATCGTGAATAATGC 704
Db 121 GCGGTGGTACTGGCTCTGGTTTATTTGTCGGAGATGGACTGATCGTGAATAATGC 180
```

```
QY 705 CCACGTGTGATCCAAACAGCAACGGGTCAAAGTTGAGTGAAGAAACGGTGCCTTACGA 764
Db 181 CCACGTGTGATCCAAACAGCAACGGGTCAAAGTTGAGTGAAGAAACGGTGCCTTACGA 240
QY 765 AGCCAAAATCAAGGATGTGGATGAGAAAACAGACATCGCACTCATCAAAAATGACCA 824
Db 241 AGCCAAAATCAAGGATGTGGATGAGAAAACAGACATCGCACTCATCAAAAATGACCA 300
QY 825 GGGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 884
Db 301 GGGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 885 GGTCCGCCATCGAAGCCCGTTTCCCTTCAAAAACAGCTCACCAACGGGATCGTGAGCAC 944
Db 361 GGTCCGCCATCGAAGCCCGTTTCCCTTCAAAAACAGCTCACCAACGGGATCGTGAGCAC 420
QY 945 CACCAGAGCGCGCAAGAGCTGGGGTCCGCAACTCAGACATGAGACTACATCCAGAC 1004
Db 421 CACCAGAGCGCGCAAGAGCTGGGGTCCGCAACTCAGACATGAGACTACATCCAGAC 480
QY 1005 CGACGCCATCATCAACTATGG--AAACTCGGAGGCGCGTTAGTAAACCTTGACGGTGA 1062
Db 481 CGACGCCATCATCAACTATGGGAACTCCGGAGGCGCGTTAGTAAACCTTGACGGTGA 540
QY 1063 GTGATTGGAATTAACACTTTGAAAGTGACAGCTGGAATCTCCTTTGCAATCCCATCTGAT 1122
Db 541 GTGATTGGAATTAACACTTTGAAAGTGACAGCTGGAATCTCCTTTGCAATCCCATCTGAT 600
QY 1123 AAGATTAAAAAGTTCTCTACGGAGTCCCATGACCGAGCGCCCAAGAGAAAGCCATCAC 1182
Db 601 AAGATTAAAAAGTTCTCTACGGAGTCCCATGACCGAGCGCCCAAGAGAAAGCCATCAC 660
QY 1183 AAGAAAGATATATTGGTATCCGAATGATGTCACTCACGTCCAGCAAGCCAAAGAGCTG 1242
Db 661 AAGAAAGATATATTGGTATCCGAATGATGTCACTCACGTCCAGCAAGCCAAAGAGCTG 720
QY 1243 AAGGACCGGCACCGGACCTTCCAGACGTGATCTCAGGAGCGTATATAATTAAGTAAT 1302
Db 721 AAGGACCGGCACCGGACCTTCCAGACGTGATCTCAGGAGCGTATATAATTAAGTAAT 780
QY 1303 CTTGATACCCAGCAGAGAGCTGT--GGTCTCAAGGAAAGAGCGTCAATATCAGATCAA 1361
Db 781 CTTGATACCCAGCAGAGAGCTGTGGGTCTCAAGGAAAGAGCGTCAATATCAGATCAA 840
QY 1362 TGGACAGTCCGTGGTCTCCGCCAATGATGTCAAGCGAG--TCATTTAAAGGAAAGCACCC 1420
Db 841 TGGACAGTCCGTGGTCTCCGCCAATGATGTCAAGCGAGTTTCAATTTAAAGGAAAGCACCC 900
QY 1421 TGAACATGTGTGCTCCAGGCGGTAAATGAAGATATCATGATCAGATGATCCGAGGAAA 1480
Db 901 TGAACATGTGTGCTCCAGGCGGTAAATGAAGATATCATGATCAGATGATCCGAGGAAA 960
QY 1481 TTGACCCCATAGGAGGATGAGCTGGAGCTTCAATGTTTCCCTCAAAGACTCTCCCGT--G 1539
Db 961 TTGACCCCATAGGAGGATGAGCTGGAGCTTCAATGTTTCCCTCAAAGACTCTCCCGTGG 1020
QY 1540 GATGACGATGAGGACTCTGGGCTGCTGGAATAGGACACTCAAGACTTTTACCTGCTT 1599
Db 1021 GATGACGATGAGGACTCTGGGCTGCTGGAATAGGACACTCAAGACTTTTACCTGCTT 1080
QY 1600 TTGTTTGTTCAGTGGAGACTCCCTGGCCCAACAGAAATCCTTCTTGATAGTTTGCAGG 1659
Db 1081 TTGTTTGTTCAGTGGAGACTCCCTGGCCCAACAGAAATCCTTCTTGATAGTTTGCAGG 1140
QY 1660 ACAAATGTAATGTGAGATCCGAGCAGAGCTCTG--CCCTCTGATCTGATGATG 1718
Db 1141 ACAAATGTAATGTGAGATCCGAGCAGAGCTCTGCCCCCTTCTGATCTGATGATG 1200
QY 1719 CAGTGTGCTTTTCTTGTCCAGCTTGGGCCAATTTCTTGCTTAGACAGTCAAGCATTTGCTCC 1778
Db 1201 CAGTGTGCTTTTCTTGTCCAGCTTGGGCCAATTTCTTGCTTAGACAGTCAAGCATTTGCTCC 1260
```



```
QY 1779 TCCTTTAACTGAGTCATCATCTTAGTCCAACTAATGAGTCAGTCGATACAAATG-CGTAGATAG 1837
Db 1261 TCCTTTAACTGAGTCATCATCTTAGTCCAACTAATGAGTCAGTCGATACAAATGCGGTAGATAG 1320
QY 1838 AAGAAGCCCAACGGAGCCAGGATGGAGCTGGTCTGCTGTTTGTGCTTTTCTCCAACTCAGC 1897
Db 1321 AAGAAGCCCAACGGAGCCAGGATGGAGCTGGTCTGCTGTTTGTGCTTTTCTCCAACTCAGC 1380
QY 1898 ACCCAAAGTCATGACACAGACCCCGGTGGGTGAGCGTGGCTTCTCAAACGCCGA 1957
Db 1381 ACCCAAAGTCATGACACAGACCCCGGTGGGTGAGCGTGGCTTCTCAAACGCCGA 1440
QY 1958 AGTTGCTCTTTTAGGAATCTTTTGAATTTGGAGCAGATGACTCTGAGTTTGAAGTA 2017
Db 1441 AGTTGCTCTTTTAGGAATCTTTTGAATTTGGAGCAGATGACTCTGAGTTTGAAGTA 1500
QY 2018 TTAAGTACTTCTTACACATTGAATAAAAAA 2051
Db 1501 TTAAGTACTTCTTACACATTGAATAAAAAA 1534

RESULT 12
AF070555
LOCUS AF070555 1423 bp mRNA linear PRI 05-AUG-1998
DEFINITION Homo sapiens clone 24795 mRNA sequence.
ACCESSION AF070555
VERSION AF070555.1 GI:3387920
KEYWORDS FLI CDNA.
SOURCE Homo sapiens
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1423)
Andersson,B., Wentland,M.A., Ricafrente,J.Y., Liu,W. and Gibbs,R.A.
A 'double adaptor' method for improved shotgun library construction
Anal. Biochem. 236 (1), 107-113 (1996)
96207227
PUBMED 8619474
REFERENCE 2 (bases 1 to 1423)
AUTHORS Yu,W., Andersson,B., Worley,K.C., Muzny,D.M., Ding,Y., Liu,W.,
Ricafrente,J.Y., Wentland,M.A., Lennson,G. and Gibbs,R.A.
Large-scale concatenation cDNA sequencing
Genome Res. 7 (4), 353-358 (1997)
9110174
REFERENCE 3 (bases 1 to 1423)
AUTHORS Yu,W. and Gibbs,R.A.
Direct Submission
TITLE Submitted (05-JUN-1998) Molecular and Human Genetics, Baylor
College of Medicine, One Baylor Plaza S930, Houston, TX 77030, USA
Location/Qualifiers
1. 1423
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="I.M.A.G.E. Consortium clone ID 24795"
/sex="female"
/tissue_type="brain"
/dev_stage="infant"
/notes="similar to Homo sapiens mRNA for serine protease
with GenBank Accession Number Y07921"

ORIGIN
Query Match 68.0%; Score 1397; DB 9; Length 1423;
Best Local Similarity 99.9%; Pred. No. 7.2e-197;
Matches 1408; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 646 CCGGTGGCTAGTGGGTCTGGGTTTATTGTGCGGAAGATGGACTGATCGTGACAAATGCC 705
Db 1 CCGGTGGCTAGTGGGTCTGGGTTTATTGTGCGGAAGATGGACTGATCGTGACAAATGCC 60
QY 706 CACGTGTGACCAACAGACCCGGGTCAAAGTTGAGTGAAGACGGTGCCTTACGAA 765
```

```
Db 61 CACGTGTGACCAACAGACCCGGTCAAAGTTGAGCTGAAGAACGGTCCCACTTACGAA 120
QY 766 GCCAAATCAAGGATGTGGATGAGAAAGCAGATCGCACTCATCAAATTTGACCCAG 825
Db 121 GCCAAATCAAGGATGTGGATGAGAAAGCAGATCGCACTCATCAAATTTGACCCAG 180
QY 826 GGCAGAGTGTCTGCTGCTTGGCGCTCTCAGAGCTCGGCGCGGAGAGTTCGTG 885
Db 181 GGCAGAGTGTCTGCTGCTTGGCGCTCTCAGAGCTCGGCGCGGAGAGTTCGTG 240
QY 886 GTGCGCATCGGAAGCCCGTTTCCCTTCAAACACAGTCACACCGGGATCGTGAAGC 945
Db 241 GTGCGCATCGGAAGCCCGTTTCCCTTCAAACACAGTCACACCGGGATCGTGAAGC 300
QY 946 ACCCAGCGAGCGGCAAGAGCTGGGCTCCGCACTCAGACATGACATACATCCAGAC 1005
Db 301 ACCCAGCGAGCGGCAAGAGCT-GGGCTCCGCACTCAGACATGACATACATCCAGAC 359
QY 1006 GAGCCATCATCAACTATGGAACCTCGGAGGCCGTTAGTAAACCTGACGCTGAAGT 1065
Db 360 GAGCCATCATCAACTATGGAACCTCGGAGGCCGTTAGTAAACCTGACGCTGAAGT 419
QY 1066 ATTGGAATTAACTTTGAAAGTGACAGCTGGAATCTCTTTGCAATCCCATCTGATAAG 1125
Db 420 ATTGGAATTAACTTTGAAAGTGACAGCTGGAATCTCTTTGCAATCCCATCTGATAAG 479
QY 1126 ATTAAGATTTCTCAGGAGTCCCATGACGACAGGCGCAAGAAAGAAAGCCATCAGC 1185
Db 480 ATTAAGATTTCTCAGGAGTCCCATGACGACAGGCGCAAGAAAGAAAGCCATCAGC 539
QY 1186 AAGAAGTATATTGGTATCCGAATGATGTCACTCAGTCCAGCAAGCCAAAGAGCTGAAG 1245
Db 540 AAGAGTATATTGGTATCCGAATGATGTCACTCAGTCCAGCAAGCCAAAGAGCTGAAG 599
QY 1246 GACCGCACCGGAGCTTCCAGACGATGTCTCAGGAGCGTATATTAATTGAAGTAACTCCT 1305
Db 600 GACCGCACCGGAGCTTCCAGACGATGTCTCAGGAGCGTATATTAATTGAAGTAACTCCT 659
QY 1306 GATACCCACAGAGAGCTGGTGTCTCAAGAAAAACAGCTCATATATCAGCATCAATGGA 1365
Db 660 GATACCCACAGAGAGCTGGTGTCTCAAGAAAAACAGCTCATATATCAGCATCAATGGA 719
QY 1366 CAGTCCGTGTCTCCGCAATGATGTCAAGACGCTATTAAAGGAAAGACCCCTGAAC 1425
Db 720 CAGTCCGTGTCTCCGCAATGATGTCAAGACGCTATTAAAGGAAAGACCCCTGAAC 779
QY 1426 ATGTGTCTCCGACGGGTAAATGAAGATATCATGATCAGATGATTCGCCAAGAAATTGAC 1485
Db 780 ATGTGTCTCCGACGGGTAAATGAAGATATCATGATCAGATGATTCGCCAAGAAATTGAC 839
QY 1486 CCATAGCAGAGGATGAGCTGGACTTCATGTTCCCTCAAAGACTCTCCCGTGGATGAC 1545
Db 840 CCATAGCAGAGGATGAGCTGGACTTCATGTTCCCTCAAAGACTCTCCCGTGGATGAC 899
QY 1546 GGATGAGACTCTGGCTGCTGGAAATAGACACTCAAGACTTTTGTACTGCAATTTGTTT 1605
Db 900 GGATGAGACTCTGGCTGCTGGAAATAGACACTCAAGACTTTTGTACTGCAATTTGTTT 959
QY 1606 GTTCAGTGGAGACTCCCTGGCCAAACAGAAATCTCTTGTAGATTTTCAGGCAAAACAAAT 1665
Db 960 GTTCAGTGGAGACTCCCTGGCCAAACAGAAATCTCTTGTAGATTTTCAGGCAAAACAAAT 1019
QY 1666 GTAATGTCAGATCCCGCAGCAGAGCTCTGCCCTCTGTATCTCTATGTATGATGAGTGTG 1725
Db 1020 GTAATGTCAGATCCCGCAGCAGAGCTCTGCCCTCTGTATCTCTATGTATGATGAGTGTG 1079
QY 1726 CTTTTCCTCCAGCTTGGGCAATCTTTGCTTTAGACAGTCAGCATTTGCTCTCTCTTTA 1785
Db 1080 CTTTTCCTCCAGCTTGGGCAATCTTTGCTTTAGACAGTCAGCATTTGCTCTCTCTTTA 1139
QY 1786 ACTGAGTCATCATCTTAGTCCAACTAATGCAAGTCGATCAATCAATGCGTAGATAGAGAGCC 1845
```


/strain="C57BL/6"
 /db_xref="taxon:10090"
 /tissue_type="ovary"
 41..1483
 /note="IGFBP-5 endopeptidase; serine protease"
 /codon_start=1
 /product="insulin-like growth factor binding protein 5
 protease"
 /protein_id="AADS2682.1"
 /db_xref="GI:5815459"
 /translation="MQSLRITLLSLLLLLAAPSLPSGTSRSPAAATVCEHCDPT
 RCAPPDPCGGRRVDCAGCEVCGALGEGPCGEGLCQVPPGVFASATVR
 RRAAGLCVCASSPVCSDAKTYNLCLEAARSSEKLRPFPVILVQRCAGGQGE
 DPNSLRHYNFADVPIKPAVVIHLYELKPLFSEVPVAGSGFVISEDGLIVTN
 AHVTKNRYKVELKNGATYAKIDVDEKADIALIKIDHOKLPVLLGRSELRPG
 EFVVAIGSPFLQNTVTGIVSTTQRCGKELGRNSDMYIOTDAIINVGNSGGLVN
 LDEVIQINTLKVTAQISPAIPSDIKKFLTESHDQAKKAVTKKXIGIRMSLTS
 SKAKELKDRHDPDLVLSGAVIIEVDPDPAEAGLKENDDVIIISNGSVVTVANDVSD
 VIKKENTLNMVVRGNEDEVITVIFPEIDP"

ORIGIN

Query Match 64.3%; Score 1321; DB 10; Length 2005;
 Best Local Similarity 82.1%; Pred. No. 1.2e-185;
 Matches 1648; Conservative 0; Mismatches 330; Indels 29; Gaps 10;

QY	29	CCGCGCGCGCAGAGTCGCATGAGATCCCGCGCGCGCTCTTCTCCGCTGCTGCTGC	88
DB	21	CGTGCGCGCTGGAGTGTGTATGAGTCCCTGCGTACACGCTCTGTCTTGTACTGC	80
QY	89	TGCTGTGGCGCGCGCGCTCGCGCAGCTGTCCCGCGCGCGCTCGCGCGCTTTGG	148
DB	81	TGCTGTAGCGCGCTCTTCTTGGCGTTCGCGTGGGACCGCGCGCTCGCGCGCGCTG	140
QY	149	CCGCGCGCGTCCAGAGACCGCTGCGAGCGCGCGCTGCGCGCGCGCGAGCACTGG	208
DB	141	CCACCGTCTGTCGAGACACTCGATCCACCGCGTGGCGCGCGCGCGCGCGCTGG	200
QY	209	AGGCGCGCGCGCGCGCGCTGCGAGCTGTGCGAGGTGTGCGCGCGCGCGCGCGCG	268
DB	201	AGGCGCGCGCGTCCGCGACCGCTGCGCGCTGTGCGAGGTGTGCGCGCGCGCTCG	260
QY	269	CCGCGTGGCGCGCTCGAGAGCGCGCTGCGCGCGCGCGCTGCGAGTGCCTGCGCT	328
DB	261	CAGCGTGGCGCGCTCGAGAGGCTCCCTGCGCGAGGCGCTGCAATGCGTAGTGCCT	320
QY	329	GGTGCGAGCGCTCGCGACCGTGTGCGCGCGCGCGCGCGCGCGCGCTGTGTGCGCG	388
DB	321	GGTGCGCGCGCTCGCGACAGTACGACGCGCGCGCGCGCGCGCGCTGTGTGCGCG	380
QY	389	GCAAGCGCGCGTGTGCGGCGAGCGAGCGCAACACCTAGCGCAACCTGTGCGAGCT	448
DB	381	GCAGCGAGCGGTGTGTGTAGCGAGCGCAAGACCTACACCAACCTGTGCGAGCTGC	440
QY	449	CCGCGCGCGCGCTCGAGAGGTGCAACCGCGCGCGCGCGCGCTGTGCGAGCGCGAG	508
DB	441	CCGCGCGCGCGCTCGAGAGGTGCGCGAGCGCGCGCGCGCGCTGTGCGAGCGCGCG	500
QY	509	CCTGCGCGCGAGGCGAGAGATCCCAACAGTTTTCGCGCATAAATATATCTTATCGCG	568
DB	501	CCTGCGCGCGAGGCGAGAGATCCCAACAGTTTTCGCGCATAAATATATCTTATCGCG	560
QY	569	ACGTGTGGAGAGATGCGCGCGCGCGGTTCATATCGAATGTTCGAGAGCTTCGT	628
DB	561	ATGTGTGGAGAGATGCGCGCGCGCGGTTCATATCGAATGTTCGAGAGCTTCGT	620
QY	629	TTTCTAAACGAGAGGTGCGGTGCTGTGGTCTGTGGTTCATGTCGGAAGATGGAC	688
DB	621	TCTCGAAGAGGAGGTGCGAGTGGCGAGTGGGTGAGGATTCATGTCGGAAGATGGAC	680
QY	689	TGATCGTGAATAATGCGCGGTGTGACCAACAGCGCGGTCAAGTGTGAGTGAAGA	748
DB	681	TGATGTGACAAATGCTCACGTGTGTCACCAACAAACCGCGTCAAGGTGTGAGTGAAGA	740

QY	749	ACGTGCGCACTTACGAGCCAAAATCAAGGATGTGGATGAGAAAGCAGACATCGCACTCA	808
DB	741	ATGAGACTACCTATGAGCCAAAATCAAGGATGTGGATGAGAAAGCGGACATTCGCGTTA	800
QY	809	TCAAAATTGACACACGAGCGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	868
DB	801	TCAAGATTGACACACGAGGAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	860
QY	869	GGCGGCGAGAGTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	928
DB	861	GACCTGAGAAATTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	920
QY	929	CCGGATCGTGAGACCAACCCAGCGAGCGGCAAAAGAGTGGGGTCCGCAACTCAGACA	988
DB	921	CTGGATCGTGAGACCAACCCAGCGAGCGGCAAAAGAGTGGGGTCCGCAACTCAGACA	980
QY	989	TGGACTACATCCAGACCGACGCTATCACTCACTATGGAACCTCGGAGGCGCGTTCGTAA	1048
DB	981	TGGACTACATCCAGACCGACGCTATCACTCACTATGGAACCTCGGAGGCGCGTTCGTAA	1040
QY	1049	ACCTGCGCGTGAAGTATGGAATTAACATTTGAAAGTACAGTGGAAATCTCTTTG	1108
DB	1041	ACCTGCGTGAAGTATGGAATTAACATTTGAAAGTACAGTGGAAATCTCTTTG	1100
QY	1109	CAATCCCATCTGATAAGATTAAAGATTTCTCACGAGTCCCATGACCGACAGCGCAAG	1168
DB	1101	CAATCCCATCTGATAAGATTAAAGATTTCTCACGAGTCCCATGACCGACAGCGCAAG	1160
QY	1169	GAAAGCGATCCAGAAAGATATTTGATTCGGAATGATGTCACCTCAGTCCAGCA	1228
DB	1161	GAAAGCGTGTCCAGAAAGATATTTGGAATGATGTCGCTCAGTCCAGCA	1220
QY	1229	AAGCAAGAGCTGAAGACCGGCAACCGGACTTCCACAGCTGATCTCAGGAGCGTATA	1288
DB	1221	AAGCAAGAGCTGAAGACCGGCAACCGGACTTCCACAGCTGATCTCAGGAGCGTATA	1280
QY	1289	TAATGAGATTTCTGATACCCAGCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1348
DB	1281	TCATTGAGTCTTCTGATACCCAGCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1340
QY	1349	TAATCAGCATCAATGGAGCAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1408
DB	1341	TCATCAGCATCAATGGAGCAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1400
QY	1409	GGGAAAGCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1468
DB	1401	AGGAAAGCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1460
QY	1469	TTCCGAGAAATTTGACCCATAGGAGGCGATGAGTGGACTTCATGTTTCCCTCAAG	1528
DB	1461	TTCCGAGAAATTTGACCCATAGGAGGCGATGAGTGGACTTCATGTTTCCCTCAAG	1520
QY	1529	ACTTCCGCTGATGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG	1588
DB	1521	A--CTCAGGGGATGGCGCATGAGAA--CTGAGCTGGTGAAGCAGACACCCGGACCTT	1576
QY	1589	TGACTGCCATTTTGTGTTTGTGAGTGGAGTCCCTGGGCAACAGAAATCTCTGTAT	1648
DB	1577	TGTCGAACATTTGCTTGTTCAGGGAACACCCCT--GCCGACAGAAATCTCTGTATG	1634
QY	1649	TTGAGCAAAACAAATGATGTTGAGATCCGAGGCGAGAGCTCTGCGCTCTCTGTAT	1708
DB	1635	TTGCGGCAAAACAAATGATGTTGAGTGGTGAAGGAGAGTGGTGGTGGTGGTGGTGGT	1693
QY	1709	CCTATGATGAGTGTGTTTCTTGTGAGTGGGCAATCTTCTGTTAGACAGTCCAGC	1768
DB	1694	GCTATGATGAGGCTGCTTCTTCTTACAGCTTGGGCTGTTCTCTGTTACACAGTCAAC	1753
QY	1769	ATTGCTCTCTCTCTTAACTGATCATCTTGTAGTCCAACTAATGAGTGGATGACATG	1828
DB	1754	ATTGCTCTCTCTCTTAACTGATCATCTTGTAGTCCAACTAATGAGTGGATGACATG	1805
QY	1829	CGTAGATAGAAAGACCCCAACCGGAGCGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG	1887

```

Db      1806 GTAGGCAAGAGAGTCTCTC-----SCAGGCAAGTCTTGGGCGTGTATTACATTTTCCT 1861
Qy      1888 CCAAGTCAGCACCCAAAGG---TCAATGCAAGAGACCCCGGTGGGTGAGCGGTGGTT 1944
Db      1862 CCGAGTCGCACTCAGAGGAATTTGATGCCCGGAGACACAGAGTGGGTGA-CCCTGGGTT 1920
Qy      1945 CTCAAACGCCCAAGTGGCTCTCTTTAGGAATCTCTTTAGGAATTTGGAGACAGATGACTC 2004
Db      1921 CTGAGATGCCCAAAATGCTCTT-----AGGAATCTGGAACCCGAGTACAGTACTT 1975

Qy      2005 TGAGTTTGGCTATTAAAGTACTTCTT 2031
Db      1976 TGAGTTTTCAGCTATTAAATACTTCTT 2002

RESULT 15
AK128916      2007 bp      mRNA      linear      ROD 09-SEP-2003
LOCUS      Mus musculus cDNA fis, clone TRACH2024749, highly similar to Serine
DEFINITION      protease HTRAI precursor (EC 3.4.21.-).
ACCESSION      AK128916
VERSION      AK128916.1 GI:34536573
KEYWORDS      oligo capping; fis (full insert sequence).
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1
Kanehori,K., Ishibashi,T., Chiba,Y., Fujimori,K., Hiraoka,S.,
Tanai,H., Watanabe,S., Ishida,S., Ono,Y., Hotuta,T., Watanabe,M.,
Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Ota,T., Wakamatsu,A.,
Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K.,
Nishikawa,T., Kimura,K., Matsu,K., Nakamura,Y., Sekine,M.,
Kikuchi,H., Kanda,K., Wadatsuma,M., Takahashi-Fujii,A., Oshima,A.,
Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K.,
Masuno,Y., Nagai,K. and Isogai,T.
NEDO cDNA sequencing project
Unpublished
2 (bases 1 to 2007)
Isogai,T. and Yamamoto,J.
Direct Submission
Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 252-0818, Japan.
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO cDNA sequencing project supported by Ministry of Economy,
Trade and Industry of Japan; cDNA full insert sequencing: Research
Association for Biotechnology (RAB); cDNA library construction:
Helix Research Institute (HRI) (supported by Japan Key Technology
Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and
Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
location/Qualifiers
1. .2007
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="TRACH2024749"
/note="cloning vector: pME18SFL3"

FEATURES
source
1. .2007

ORIGIN
Query Match      64.3%; Score 1321; DB 10; Length 2007;
Best Local Similarity 82.1%; Pred. No. 1.2e-185;
Matches 1648; Conservative 0; Mismatches 330; Indels 29; Gaps 10;

Qy      29 CGCGCGCCGACAGTGGCGATCGATCCGATCCGCGCGCGTCTCTCCGCGTGTGCTGC 88
Db      23 CGGTGCCGTCGAGTGGTATGATCGATCGTCCCTGCGTACACGCTCTCTTTTGTACTGCG 82
Qy      89 TGCTGTGCGCGCGCGCTCTCGCGCAGTGTCCCGCGCGCGCGCTCGCGCGCTTTGG 148
Db      83 TSGTGTACGCGCTCTCTCTTGGCGTGTGCGTGGGACCGGCGCGCTCGGCGCGAGCTG 142

```

```

Qy      149 CGCGCGGTGCCAGACCGCTGGAGCGCGGCGCGCTGCCCGCGCAGCGCGAGCACTGCG 208
Db      143 CCACCGTCTGTCCGAGCACTGGGATCCACCCCGCTGGCGCGCGCGCCACCGGACTGCG 202
Qy      209 AGGCGCGCGCGCGCGCGCGTGGCGGACCGTGGCGGTGCTGCGAGGTGTGCGCGCGCGCGCGGCG 268
Db      203 AGGCGCGCGCGTCCGCGACCGTGGCGTGTGCGAGGTGTGCGCGCGCGCTCGAGGGTG 262
Qy      269 CGCGTGGCGCTCGAGGAGGCGCGTGGCGGAGGCGCTCCAGTGGTGGTGGTGGTGGTGGTGGTGG 328
Db      263 CAGCGTGGCGCTCGAGGAGGCGTCCCTGGCGGAGGCGCTGCAATGCGTAGTGGCGCTTGG 322
Qy      329 GGTGCGCAGCTCGGCCACCGTGGCGGCGCGCGCGCGCGCGCTCTGTGTGTGCGCCA 388
Db      323 GGTGCGCGCTCGGCCACAGTACGACGCGCGCGCACAGCGCGCGCTGTGCGTGTGTGCCA 382
Qy      389 GCAGCAGCGCGTGTGGCGCAGCGACGCCACACTACCGCAACCTGTGCCAGCTGCGCG 448
Db      383 GCAGCAGCGCGTGTGGTAGCGACCGCAAGACTCACCAACCTGTGCCAGCTGCGCG 442
Qy      449 CGCCAGCGCGCTCGAGAGGTGCACCGCGCGCGGTGCATCGTCTCTGCGAGCGGAG 508
Db      443 CGCCAGCGCGCTCGGAGAGCTTCGCGAGCGCGCGGTGCATCGTCTGCGAGCGCGG 502
Qy      509 CTTGCGCGCAAGGCGAGGATCCCAACAGTTTGGCGCCATAAATATACTTTATCGCG 568
Db      503 CTTGCGCGCAGGCGAGGATCCCAACAGTTTGGCGCCATAAATATACTTTATCGCG 562
Qy      569 AGTGTGTGAGAAAGTGGCGCGTGGTTCATATCGAATTTGTTTCCGAAGTTCCT 628
Db      563 ATGTGTGTGAGAAAGTGGCGCGTGGTTCATATTTGAATTTGTTTCCGAAGTTCCT 622
Qy      629 TTTCTAAACGAGAGTGGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 688
Db      623 TCTCGAGAGGAGGTGGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 682
Qy      689 TGATCGTGAACAATGCCAGTGGTGCACCAACAGCCGCGTCAAAGTTGAGTGAAGA 748
Db      683 TGATGTGAACAATGCTACGTGGTGCACCAACAAACCCGGTCAAGGTGAGTGAAGA 742
Qy      749 AGGTGTCCTTACGAGCCAAATCAAGATGGATGAGAAAGCAGACATCGCACTCA 808
Db      743 ATGAGCTTACCTATGAGCCAAATCAAGATGGATGAGAAAGCAGACATCGCACTTA 802
Qy      809 TCAAAATTGACCAACCGGCAAGTGCCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 868
Db      803 TCAAGATTGACCAACCGGAAAGCTGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 862
Qy      869 GCGCGGAGAGTTCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 928
Db      863 GACCTGGAGAAATTTGATTTGCCATTGGAAGCCCTTTCTTTCAAACACAGTCAACA 922
Qy      929 CGCGGATCGTGAGCACCAACCGCGCGCGGCAAGAGTGGGCTCCGCAACTCAGACA 988
Db      923 CTGGGATCGTGAGCACCAACCGCGCGCGGCAAGAGTGGGACTTCGGAATCCGATA 982
Qy      989 TGGCTACATCCAGCGCGCGCATCATCACTATGAACTCGGAGGCGCGTGTAGTAA 1048
Db      983 TGGACTACATCCAGCAGCGCTATCATCAATTAAGAAATTCGAGGCGCGTGTAGTAA 1042
Qy      1049 ACCTGGACGCTGAAGTGAATTAACACTTTGAAAGTGACAGCTGGAATCTCTTTG 1108
Db      1043 ACCTGGATGCGAGGTGATTGGGATTAAACACTTTGAAGTGACGCGCGGCTCTCTTCG 1102
Qy      1109 CAATCCCATCTGATAGATTAAAGTTCCTCAGGAGTCCCATGACCGCAGGCGCAAG 1168
Db      1103 CAATTCATCCGATAAGATAAAAGTTCTTGACAGAGTCCCATGATCGACAGGCGCAAG 1162
Qy      1169 GAAAGCCATCAACCAAGAAAGTATATTGGTATCCGAATGATGTCACTCAAGTCCAGCA 1228
Db      1163 GGAAGCTGTCCCAAGAAAGATATTATGGGATCCGAATGATGTGCTCATATCTAGCA 1222

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

M nucleic - nucleic search, using sw model

Run on: March 10, 2004, 08:01:57 ; Search time 499.903 Seconds

(without alignments)

17454.974 Million cell updates/sec

Title: US-10-084-817-32

Perfect score: 2054

Sequence: 1 cgggctccgtccgtccgc.....cattgaaaaaaaaaaaaa 2054

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: N_Geneseq_29Jan04.*
- 2: Geneseq1980s.*
- 3: Geneseq1990s.*
- 4: Geneseq2000s.*
- 5: Geneseq2001as.*
- 6: Geneseq2001bs.*
- 7: Geneseq2002as.*
- 8: Geneseq2003as.*
- 9: Geneseq2003bs.*
- 10: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2037.4	99.2	2157	6	ABQ54255 Human ova
2	2034.4	99.0	2036	6	ABV78051 Hypoxia-x
3	2034.4	99.0	2075	2	AAT75444 cDNA enco
4	2034.4	99.0	2036	2	AAT75444 Homo sapi
5	2010.8	97.9	2205	2	AAV29540 Human pre
6	1780	86.7	1855	3	AAF21747 Human bre
7	1701.4	82.8	1894	9	ADB62641 Human cDN
8	1644	80.0	1705	2	Aaz41293 Human nor
9	1455.2	70.8	1534	4	Aah23090 Osteoarth
10	1440	70.1	1440	6	AD44188 Human TAN
11	1191.8	58.0	1443	3	Aaz29176 Recombina
12	1043.8	50.8	1185	3	Aaz29195 cDNA enco
13	759.8	37.0	1389	7	ABX77560 Different
14	635	30.9	669	6	ABQ59513 Human col
15	558.4	27.2	900	4	AAL24929 Human bre
16	530	25.8	895	4	AAL24918 Human bre
17	493.4	24.0	495	6	ABK38859 cDNA enco
18	493.4	24.0	495	7	ACAL1188 Human lun
19	493.4	24.0	495	7	ACAL1188 Lung can
20	478.2	23.3	487	3	AAA30449 Human bra
21	451.6	22.0	1797	6	AAD23855 Human pro
22	451.6	22.0	2094	6	ABK51497 cDNA enco
23	451.6	22.0	2543	7	ACF12810 Human PRS

24	451.6	22.0	2577	3	AAA57359
25	447.2	21.8	607	2	AXX29130
26	447.2	21.8	1583	7	ACF16993
27	446	21.7	490	8	ACH26037
28	427.4	20.8	2450	7	ACF12809
29	414.6	20.2	1544	4	AAK94547
30	414.6	20.2	1553	2	AAV43618
31	406.6	19.8	2040	3	Az52362 NSEQ gene
32	406.6	19.8	2040	6	AAI42458
33	402.2	19.6	457	8	ACH13892
34	395.8	19.3	1044	6	AD44189
35	395.8	19.3	1563	3	AAA57361
36	391.4	19.1	1953	7	ACF12811
37	389	18.9	705	4	AAI24879
38	386.4	18.8	1695	4	AAI4885
39	385.6	18.8	1788	4	AAI26848
40	384	18.7	384	7	ACF12817
41	381.8	18.6	1859	4	AAI10126
42	381.8	18.6	1859	6	ABL35059
43	365.6	17.8	1897	7	ACF12815
44	352.8	17.2	412	7	ABX39463
45	331.8	16.2	1193	6	ABQ30909

ALIGNMENTS

RESULT 1

ABQ54255

ID ABQ54255 standard; cDNA; 2157 BP.

XX AC ABQ54255;

XX AC ABQ54255;

DT 22-AUG-2002 (first entry)

XX DE Human ovarian antigen HKADF15 cDNA, SEQ ID NO:135.

Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
ovarian cancer; breast cancer; tumour; reproductive system disorder;
infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
inflammatory condition; immune disorder; blood disorder;
cardiovascular disorder; respiratory disorder; neurological disorder;
gastrointestinal disorder; urinary system disorder; drug screening;
gene therapy; chromosome mapping; forensic analysis;
antibody preparation; cytostatic; immunomodulatory; neuroprotective;
antiinflammatory; gynaecological; reproductive; chromosome 10q25.3-26.2;
gene; ss.

XX OS Homo sapiens.

XX PN WO200200677-A1.

XX PD 03-JAN-2002.

XX PF 07-JUN-2001; 2001WO-US018569.

XX PR 07-JUN-2000; 2000US-0209467P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Birse CE, Rosen CA;

XX DR WPI; 2002-147878/19.

XX DR P-ESDB; ABP41178.

Isolated nucleic acid molecules encoding novel ovarian polypeptides,
useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
cancer), immune disorders, cardiovascular disorders and neurological
diseases.

Claim 1; SEQ ID NO 135; 2922pp; English.

793	CGGGAGCCCGCTTAGTAAACCTGGACCGGTGAAGTGGATTTGGAATTAACACATTTTGAAAGTGA	850
1091	CAGCTGGAATCTCCTTTTGCAATCCCATCTGATTAAGATTAATAAGTTCTCTCACGGAGTCCC	1150
853	CAGCTGGAATCTCCTTTTGCAATCCCATCTGATTAAGATTAATAAGTTCTCTCACGGAGTCCC	912
1151	ATGACCGACAGGCGCAAGGAAAGCCATCACCAAGAGAAAGTATATTGGTATCCGAATGA	1210
913	ATGACCGACAGGCGCAAGGAAAGCCATCACCAAGAGAAAGTATATTGGTATCCGAATGA	972
1211	TGTCACTACGTCCTCAGCAAGCCCAAGAGCTGAAGGACCGGCACCGGGACTTCCACAGCG	1270
973	TGTCACTACGTCCTCAGCAAGCCCAAGAGCTGAAGGACCGGCACCGGGACTTCCACAGCG	1032
1271	TGATCTCAGGAGCGGTATATAATTTGAAGTAATTTCTTGATACCCACAGCAGAAGCTGGTGGTC	1330
1033	TGATCTCAGGAGCGGTATATAATTTGAAGTAATTTCTTGATACCCACAGCAGAAGCTGGTGGTC	1092
1331	TCAAGGAAACAGCGCTCATTAATCAGCATCAATGGACAGTCCGTGGTCTCGGCCAATGATG	1390
1093	TCAAGGAAACAGCGCTCATTAATCAGCATCAATGGACAGTCCGTGGTCTCGGCCAATGATG	1152
1391	TCAGCGAGCTCNTTAAAGGGAAGCACCTGACATGGTGGTCCGACAGGGGTAAATCAAG	1450
1153	TCAGCGAGCTCNTTAAAGGGAAGCACCTGACATGGTGGTCCGCA - GGGTAATGAAG	1211
1451	ATATCATGATCACAGTGATTTCCGGAAGAAATGACCCATAGGCAGAGGCGATAGCTGGAC	1510
1212	ATATCATGATCACAGTGATTTCCGGAAGAAATGACCCATAGGCAGAGGCGATAGCTGGAC	1271
1511	TTCAATGTTTCCCTCAAGACTCTCCCGTGATGACGGATGAGGACTCTGGGCTCTCGGAA	1570
1272	TTCAATGTTTCCCTCAAGACTCTCCCGTGATGACGGATGAGGACTCTGGGCTCTCGGAA	1331
1571	TAGGACACTCAAGACTTTTGACTGCCATTTTGTGTTGTTTCAGTGAGAGACTCCCTGGCCAC	1630
1332	TAGGACACTCAAGACTTTTGACTGCCATTTTGTGTTGTTTCAGTGAGAGACTCCCTGGCCAC	1391
1631	AGAATCCTTCTTGATAGTTTGCAGGCCAAAAACAATGTAATGTTTGCAGATCCGCAGGCAGA	1690
1392	AGAATCCTTCTTGATAGTTTGCAGGCCAAAAACAATGTAATGTTTGCAGATCCGCAGGCAGA	1451
1691	AGCTCTGCCCTTCTGTAATCTATGTAATGACGTGTCTTTTCTTGCCAGACTTTGGGCCAAT	1750
1452	AGCTCTGCCCTTCTGTAATCTATGTAATGACGTGTCTTTTCTTGCCAGACTTTGGGCCAAT	1511
1751	CTTGCTTTAGACAGTCAAGTATTTGTCCTCTCTTTAACTGAGTCATCATCTTAGTCCAAC	1810
1512	CTTGCTTTAGACAGTCAAGTATTTGTCCTCTCTTTAACTGAGTCATCATCTTAGTCCAAC	1571
1811	AATGACGTTCAGATCAATTCGCTAGATGAAGAGGCCCAAGCCACCGGAGCCAGGATGGACTGGT	1870
1572	AATGACGTTCAGATCAATTCGCTAGATGAAGAGGCCCAAGCCACCGGAGCCAGGATGGACTGGT	1631
1871	CGTGTTTGTGTTTTCTTCCAGTTCAGCACCCCAAGAGTCAATGACACAGAGACCCCGGGTGG	1930
1632	CGTGTTTGTGTTTTCTTCCAGTTCAGCACCCCAAGAGTCAATGACACAGAGACCCCGGGTGG	1691
1931	GTGAGCGCTGGCTTCTCAAACGGCCGAAGTTGCCCTCTTTTAGGAAATCTCTTTTGGAAATGG	1990
1692	GTGAGCGCTGGCTTCTCAAACGGCCGAAGTTGCCCTCTTTTAGGAAATCTCTTTTGGAAATGG	1751
1991	GAGCACGATGACTCTGAGTTTGACTATTAAGTACTCTTTACACATTTGAACAAAAAAA	2050
1752	GAGCACGATGACTCTGAGTTTGACTATTAAGTACTCTTTACACATTTGAACAAAAAAA	1811
2051	AAAA 2054	
1812	AAAA 1815	

QY 1031 CGGGAGGCCCGTTAGTAAACCTGGACGGTGAAGTGATTGGAATTAAACACATTGAAACTGA 1090

QY 1309 ACCGACGAGAGCTGGTCTCAAGGAAAAAGACGTCATAATCAGCATCATGACAG 1368
DB |||||
QY 1164 ACCGACGAGAGCTGGTCTCAAGGAAAAAGACGTCATAATCAGCATCATGACAG 1223
DB |||||
QY 1369 TCCGTGGTCTCCGCAATGATGTGAGGACGTCATTAAAGGAAAGACCCCTGAACTG 1428
DB |||||
QY 1224 TCCGTGGTCTCCGCAATGATGTGAGGACGTCATTAAAGGAAAGACCCCTGAACTG 1283
DB |||||
QY 1429 GTGGTCCGAGGGGTAATGAGATATCATGATCATGATGATCCCGAAGAAATGACCCA 1488
DB |||||
QY 1284 GTGGTCCGAGGGGTAATGAGATATCATGATCATGATGATCCCGAAGAAATGACCCA 1343
DB |||||
QY 1489 TAGGACGAGCATGAGCTGGACTTCATGTTTCCCTCAAAGACTCTCCCGTGGATGACGGA 1548
DB |||||
QY 1344 TAGGACGAGCATGAGCTGGACTTCATGTTTCCCTCAAAGACTCTCCCGTGGATGACGGA 1403
DB |||||
QY 1549 TGAGGACTCTGGCTGCTGGAATAGGACACTCAAGACTTTTGACTGCCATTTGTTGTT 1608
DB |||||
QY 1404 TGAGGACTCTGGCTGCTGGAATAGGACACTCAAGACTTTTGACTGCCATTTGTTGTT 1463
DB |||||
QY 1609 CAGTGAGAGCTCCCTGGCCCAACAGAACTCTTTGATAGTTTGCAAGGCAAAACAAATGTA 1668
DB |||||
QY 1464 CAGTGAGAGCTCCCTGGCCCAACAGAACTCTTTGATAGTTTGCAAGGCAAAACAAATGTA 1523
DB |||||
QY 1669 ATGTTGCAGATCCGACGACAGAGCTCTGCCCTTCTGTATCTATGATGAGTGTGCTT 1728
DB |||||
QY 1524 ATGTTGCAGATCCGACGACAGAGCTCTGCCCTTCTGTATCTATGATGAGTGTGCTT 1583
DB |||||
QY 1729 TTTCTTCCAGCTTGGGCCATTCTTGCTTAGACAGTCAGCATTTGTCTCTCTCTTAAC 1788
DB |||||
QY 1584 TTTCTTCCAGCTTGGGCCATTCTTGCTTAGACAGTCAGCATTTGTCTCTCTCTTAAC 1643
DB |||||
QY 1789 GAGTCATCATCTTAGTCCAACTAATGACGTGATACAAATGCGTAGATAGAGAACCCCA 1848
DB |||||
QY 1644 GAGTCATCATCTTAGTCCAACTAATGACGTGATACAAATGCGTAGATAGAGAACCCCA 1703
DB |||||
QY 1849 CGGGAGCCAGATGGGACTGGTCTGTTTGTGCTTTTCTCCAAGTCAGCACCCCAAGGTC 1908
DB |||||
QY 1704 CGGGAGCCAGATGGGACTGGTCTGTTTGTGCTTTTCTCCAAGTCAGCACCCCAAGGTC 1763
DB |||||
QY 1909 AATGCACAGACCCCGGTGGGTGAGCGCTGGCTTCTCAAACGCGCAAGTGTGCTCTT 1968
DB |||||
QY 1764 AATGCACAGACCCCGGTGGGTGAGCGCTGGCTTCTCAAACGCGCAAGTGTGCTCTT 1823
DB |||||
QY 1969 TTAGGAATCTCTTTGGAATTTGGGAGCAGATGACTCTGAGTTTGAGCTATTAAAGTACTT 2028
DB |||||
QY 1824 TTAGGAATCTCTTTGGAATTTGGGAGCAGATGACTCTGAGTTTGAGCTATTAAAGTACTT 1883
DB |||||
QY 2029 CTTACACATTG 2039
DB |||||
QY 1884 CTTACACATTG 1894
DB |||||

RESULT 8
AAZ41293
ID AAZ41293 standard; cDNA; 1705 BP.

AC AAZ41293;

DT 18-JAN-2000 (first entry)

DE Human normal ovarian tissue derived cDNA 72.

XX Human; ovary; screening; ovarian cancer; treatment; ss.

OS Homo sapiens.

XX DE19816395-Al.

PD 07-OCT-1999.

XX 03-APR-1998; 98DE-01016395.

XX 03-APR-1998; 98DE-01016395.
XX (META-) METAGEN GES GENOMFORSCHUNG MBH.
XX Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
XX WPI; 1999-552352/47.
XX Nucleic acid sequences potentially useful in diagnosis or therapy of
XX ovarian cancer.
XX Claim 3; Page 182; 274pp; German.
XX This invention describes novel nucleic acid sequences that are highly
XX expressed in normal ovary tissue. Artificial chromosomes and cosmid
XX clones containing the sequences can be used as gene transfer vehicles.
XX The sequences can be used to produce DNA fragments containing full-length
XX genes. Host cells transformed with the sequences can be used to produce
XX polypeptides or polypeptide fragments, which can be used to screen phage
XX displays for polypeptides that bind to them, or as tools for identifying
XX agents active against ovarian cancer, or to prepare medicaments for
XX treating ovarian cancer. The cDNA sequences can be used to obtain genomic
XX genes, their promoters, enhancers, silencers, exon structures, intron
XX structures and their splice variants. AAZ4122-241324 represent cDNA
XX sequences derived from normal human ovarian tissue and which encode the
XX protein fragments represented in AAY59724-Y59837
SQ Sequence 1705 BP; 469 A; 426 C; 444 G; 366 T; 0 U; 0 Other;

Query Match 80.0%; Score 1644; DB 2; Length 1705;
Best Local Similarity 99.9%; Pred. No. 1.7e-312;
Matches 1666; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 387 CAGCAGCGAGCCGGTGTGCGCAGCGACGCCAACACCTACGCCAACCTGTGCCAGTGC 446
DB 1 CAGCAGCGAGCCGGTGTGCGCAG - GACGCCAACACCTACGCCAACCTGTGCCAGTGC 59
QY 447 CGCGCCAGCCCGCTCCGAGAGGCTGCACGCCGCCGCGTCATCGTCTCGACGCGCG 506
DB 60 CGCGCCAGCCCGCTCCGAGAGGCTGCACGCCGCCGCGTCATCGTCTCGACGCGCG 119
QY 507 AGCTGCGGCGAAGGCGAGGAAGATCCCAAGTTTGGGCCATAATAAATTAACCTTTATCG 566
DB 120 AGCTGCGGCGAAGGCGAGGAAGATCCCAAGTTTGGGCCATAATAAATTAACCTTTATCG 179
QY 567 GGACGTGGTGGAGAAGATCGCCCTGCGTGTTCATATCGAATTTGTTTCGCAAGCTTCC 626
DB 180 GGACGTGGTGGAGAAGATCGCCCTGCGTGTTCATATCGAATTTGTTTCGCAAGCTTCC 239
QY 627 GTTTTCTAAACGAGAGGTGCGGTGGCTAGTGGGTCTGGGTATTATGTCTCGAAGATGG 686
DB 240 GTTTTCTAAACGAGAGGTGCGGTGGCTAGTGGGTCTGGGTATTATGTCTCGAAGATGG 299
QY 687 ACTGATCTGACAAATGCCACGCTGGTGACCAACAAAGCACCGGGTCAAAAGTTGAGCTGAA 746
DB 300 ACTGATCTGACAAATGCCACGCTGGTGACCAACAAAGCACCGGGTCAAAAGTTGAGCTGAA 359
QY 747 GAACGCTGCCACTTACGAAGCCAAAATCAAGGATGTGATGAGAAAGCAGACATCGCACT 806
DB 360 GAACGCTGCCACTTACGAAGCCAAAATCAAGGATGTGATGAGAAAGCAGACATCGCACT 419
QY 807 CATCAAAATTTGACCAACGAGGCAAGCTGCTGTCTGTGTGCTGCTGCTGCTGCTGCTGCT 866
DB 420 CATCAAAATTTGACCAACGAGGCAAGCTGCTGTCTGTGTGCTGCTGCTGCTGCTGCTGCT 479
QY 867 GCGGCGGGGAGAGTTGCTGTGCGCATCGGAAGCCCGCTTTTCCCTTCAAAAACAGTCA 926
DB 480 GCGGCGGGGAGAGTTGCTGTGCGCATCGGAAGCCCGCTTTTCCCTTCAAAAACAGTCA 539
QY 927 CACCGGATCGTGAGCACCAACCGAGGCGGCAAGAGCTGGGGCTCCGCAACTCAGA 986
DB 540 CACCGGATCGTGAGCACCAACCGAGGCGGCAAGAGCTGGGGCTCCGCAACTCAGA 599

QY 987 CATGACTACATCCAGACCGCCATCATCAACTATGAAACTCGGAGGCCCGTTAGT 1046
 Db 600 CATGACTACATCCAGACCGCCATCATCAACTATGAAACTCGGAGGCCCGTTAGT 659
 QY 1047 AAACCTGGACCGTGAAGTATGGAATTAACACTTTGAAGTGACAGCTGGATCTCCCT 1106
 Db 560 AAACCTGGACCGTGAAGTATGGAATTAACACTTTGAAGTGACAGCTGGATCTCCCT 719
 QY 1107 TGCAATCCATCTGATAGATTAAGAGTTCTCCAGGAGTCCCATGACGACAGCCAA 1166
 Db 720 TGCAATCCATCTGATAGATTAAGAGTTCTCCAGGAGTCCCATGACGACAGCCAA 779
 QY 1167 AGGAAAGCCATCACCAGAGAGAGTATATGATATCCGAATGATCTCACTACGTCACG 1226
 Db 780 AGGAAAGCCATCACCAGAGAGAGTATATGATATCCGAATGATCTCACTACGTCACG 839
 QY 1227 CAAGCCAAAGAGCTGAAGAGCCGACCGGACCTTCCAGAGCTGATCTCAGGAGCGTA 1286
 Db 840 CAAGCCAAAGAGCTGAAGAGCCGACCGGACCTTCCAGAGCTGATCTCAGGAGCGTA 899
 QY 1287 TATAATTGAAGTAACTCTGATACCCAGCAGAGCTGGTCTCAAGGAAACGACGT 1346
 Db 900 TATAATTGAAGTAACTCTGATACCCAGCAGAGCTGGTCTCAAGGAAACGACGT 959
 QY 1347 CATATCAGCATCAATGACAGCTCGTGTCTCCGCCAATGATGTCAGGAGCTCATTA 1406
 Db 960 CATATCAGCATCAATGACAGCTCGTGTCTCCGCCAATGATGTCAGGAGCTCATTA 1019
 QY 1407 AAGGAAAGCACCCTGAACATGGTCTCGCAGGAGTGAATGAAGATATCATGATCACAGT 1466
 Db 1020 AAGGAAAGCACCCTGAACATGGTCTCGCAGGAGTGAATGAAGATATCATGATCACAGT 1078
 QY 1467 GATTCCGAGAAATGACCCATAGGAGAGGATGAGCTGGAATCATGTTTCCCTCAA 1526
 Db 1079 GATTCCGAGAAATGACCCATAGGAGAGGATGAGCTGGAATCATGTTTCCCTCAA 1138
 QY 1527 AGACTCTCCCTGATGATGAGGAGTCTGGGTCTGGAATAGGACACTCAAGACT 1586
 Db 1139 AGACTCTCCCTGATGATGAGGAGTCTGGGTCTGGAATAGGACACTCAAGACT 1198
 QY 1587 TTGACTGCCATTTGTTTTCAGTGAGAGTCCCTGCGCCAAAGAACTCTTCTTGATA 1646
 Db 1199 TTGACTGCCATTTGTTTTCAGTGAGAGTCCCTGCGCCAAAGAACTCTTCTTGATA 1258
 QY 1647 GTTTCAGGCAAAACAAATGATGTTGAGATCCGAGGAGAGCTCTGCCCTTCTGT 1706
 Db 1259 GTTTCAGGCAAAACAAATGATGTTGAGATCCGAGGAGAGCTCTGCCCTTCTGT 1318
 QY 1707 ATCTATGATGCAAGTGTCTTTTCTGCGAGCTTGGCCCATTTCTGCTTAGACAGTCA 1766
 Db 1319 ATCTATGATGCAAGTGTCTTTTCTGCGAGCTTGGCCCATTTCTGCTTAGACAGTCA 1378
 QY 1767 GCATTTGCTCTCTCTTAACTGAGTCACTATCTAGTCCCACTAATGAGTGCATACAA 1826
 Db 1379 GCATTTGCTCTCTCTTAACTGAGTCACTATCTAGTCCCACTAATGAGTGCATACAA 1438
 QY 1827 TGCGTAGATAGAAGAGCCCAACGAGGACGAGATGGGACTGGTCTGTTGCTTTTC 1886
 Db 1439 TGCGTAGATAGAAGAGCCCAACGAGGACGAGATGGGACTGGTCTGTTGCTTTTC 1498
 QY 1887 TCCAAAGTCAGACCCAAAGGTCATGACAGAGACCCCGGTGGGTGAGCGCTGCTTCT 1946
 Db 1499 TCCAAAGTCAGACCCAAAGGTCATGACAGAGACCCCGGTGGGTGAGCGCTGCTTCT 1558
 QY 1947 CAACGCGCGAAGTTGCTCTTTTAGGAATCTCTTTGGAAATGGGAGCAGATGACTCTG 2006
 Db 1559 CAACGCGCGAAGTTGCTCTTTTAGGAATCTCTTTGGAAATGGGAGCAGATGACTCTG 1618
 QY 2007 AGTTTGAGCTATTAAGTACTTCTTACACATTGAAAAAATAAAAAA 2054
 Db 1619 AGTTTGAGCTATTAAGTACTTCTTACACATTGAAAAAATAAAAAA 1666

RESULT 9
 AAH23090
 ID AAH23090 standard; DNA; 1534 BP.
 XX AAH23090;
 AC AAH23090;
 XX 17-SEP-2001 (first entry)
 DT 17-SEP-2001 (first entry)
 XX Osteoarthritis tissue-derived nucleic acid sequence #20.
 DE Osteoarthritis; infectious disorder; Crohn's disease; sepsis; human;
 XX wound healing; osteopathic; anti-arthritis; anti-inflammatory; vulnery;
 KW antibacterial; antiallergic; ds.
 KW Homo sapiens.
 OS Homo sapiens.
 XX WO200153531-A2.
 FN 26-JUL-2001.
 XX 18-JAN-2001; 2001WO-US000016.
 PF 18-JAN-2001; 2000US-0176523P.
 XX 18-JAN-2001; 2000US-0176523P.
 PR (PHAA) PHARMACIA CORP.
 XX Phippard D, Vasanthakamur G, Dotson S, Ma X;
 XX WPI; 2001-451914/48.
 XX Substantially purified protein, polypeptide or their fragments, used to
 PT identify a biologically active compound or composition and treat
 PT mammalian osteoarthritis.
 XX Claim 1; Page 106-107; 144pp; English.
 XX Sequences AAH23071-23152 represent nucleic acid sequences derived from
 CC osteoarthritis tissues. The sequences are useful as probes and for the
 CC diagnosis or prognosis of mammalian osteoarthritis. The polynucleotides
 CC and polypeptides of the invention are useful for generating diagnostic
 CC reagents, as targets for small molecule drug development, generation of
 CC therapeutics, and cloning genes. Specific antibodies are used to generate
 CC enzyme linked immunosorbent assays for detection of osteoarthritis. The
 CC invented molecules can be used to treat osteoarthritis or to analyse the
 CC disease-modifying activity of osteoarthritis drugs. Other disorders
 CC treatable using the nucleic acid sequences include atopic, inflammatory
 CC and infectious disorders e.g. Crohn's disease and sepsis, and wound
 CC healing
 XX SQ Sequence-1534 BP; 416 A; 372 C; 392 G; 354 T; 0 U; 0 Other;
 Query Match 70.8%; Score 1455.2; DB 4; Length 1534;
 Best Local Similarity 99.3%; Pred. No. 1.6e-275;
 Matches 1524; Conservative 0; Mismatches 3; Indels 7; Gaps 6;
 QY 525 GGAAGATCCCAACAGTTTTCGCGCATATAATATTAATTTATCGCGAGCTGTGGAGAGAT 584
 Db 1 GGAAGATCCCAACAGTTTTCGCGCATATAATATTAATTTATCGCGAGCTGTGGAGAGAT 60
 QY 585 CGCCCTCGCGTGTTCATATCGAATTTGTCGAAGCTTCCGTTTCTAAACGAGAGGT 644
 Db 61 CGCCCTCGCGTGTTCATATCGAATTTGTCGAAGCTTCCGTTTCTAAACGAGAGGT 120
 QY 645 GCGGTGCTAGTGGTCTGGGTTTATTTGTCGGAAGATGGACTGATCGTGCATATGC 704
 Db 121 GCGGTGCTAGTGGTCTGGGTTTATTTGTCGGAAGATGGACTGATCGTGCATATGC 180
 QY 705 CCACGTGTGACCAACAACAGCAGCGGTCAAGTTGAGCTGAAGAACCGTGCCTTACGA 764
 Db 181 CCACGTGTGACCAACAACAGCAGCGGTCAAGTTGAGCTGAAGAACCGTGCCTTACGA 240
 QY 765 AGCCAAATCAAGGATGTGGATGAGAAACGACNATCGGACTCATCAAAATTGACCACCA 824


```
FT      old_sequence      /*tag= c      replace(12, G)
FT      old_sequence      /*tag= d      replace(12, G)
FT      old_sequence      /*tag= e      replace(13, C)
FT      old_sequence      /*tag= f      replace(15, C)
FT      old_sequence      /*tag= g      replace(18, C)
FT      old_sequence      /*tag= h      replace(21, T)
FT      old_sequence      /*tag= i      replace(22, C)
FT      old_sequence      /*tag= j      replace(24, T)
FT      old_sequence      /*tag= k      replace(25, C)
FT      old_sequence      /*tag= l      replace(27, C)
FT      old_sequence      /*tag= m      replace(30, G)
FT      old_sequence      /*tag= n      replace(33, G)
FT      old_sequence      /*tag= o      replace(37, C)
FT      old_sequence      /*tag= p      replace(39, G)
FT      old_sequence      /*tag= q      replace(40, C)
FT      old_sequence      /*tag= r      replace(42, G)
FT      old_sequence      /*tag= s      replace(48, G)
FT      old_sequence      /*tag= t      replace(51, G)
FT      old_sequence      /*tag= u      replace(54, G)
FT      old_sequence      /*tag= v      replace(57, C)
FT      old_sequence      /*tag= w      replace(60, C)
FT      old_sequence      /*tag= x      replace(63, G)
FT      old_sequence      /*tag= y      replace(66, G)
FT      old_sequence      /*tag= z      replace(69, G)
FT      old_sequence      /*tag= aa     replace(72, G)
FT      old_sequence      /*tag= ab     replace(75, C)
FT      old_sequence      /*tag= ac     replace(78, G)
FT      old_sequence      /*tag= ad     replace(81, C)
FT      old_sequence      /*tag= ae     replace(84, C)
FT      old_sequence      /*tag= af     replace(85, C)
FT      old_sequence      /*tag= ag     replace(87, C)
FT      old_sequence      /*tag= ah     replace(90, G)
FT      old_sequence      /*tag= ai     replace(93, G)
FT      old_sequence      /*tag= aj     replace(96, T)
FT      old_sequence      /*tag= ak     replace(99, G)
FT      old_sequence      /*tag= al     replace(102, C)
FT      old_sequence      /*tag= am     replace(105, C)
FT      old_sequence      /*tag= an     replace(108, G)
FT      old_sequence      /*tag= ao     replace(111, C)
FT      old_sequence      /*tag= ap     replace(114, A)
FT      old_sequence      /*tag= aq     replace(117, C)
FT      old_sequence      /*tag= ar     replace(118, C)
FT      old_sequence      /*tag= as     replace(120, A)
FT      old_sequence      /*tag= at     replace(123, C)
FT      old_sequence      /*tag= au     replace(129, G)
FT      old_sequence      /*tag= av     replace(132, G)
FT      old_sequence      /*tag= aw     replace(133, C)
FT      old_sequence      /*tag= ax     replace(135, C)
FT      old_sequence      /*tag= ay     replace(138, C)
FT      old_sequence      /*tag= az     replace(141, G)
FT      old_sequence      /*tag= ba     replace(144, G)
FT      old_sequence      /*tag= bb     replace(147, G)
FT      old_sequence      /*tag= bc     replace(150, G)
FT      old_sequence      /*tag= bd     replace(153, G)
FT      old_sequence      /*tag= be     replace(156, C)
FT      old_sequence      /*tag= bf     replace(162, G)
FT      old_sequence      /*tag= bg     replace(165, C)
FT      old_sequence      /*tag= bh     replace(168, C)
FT      old_sequence      /*tag= bi     replace(169, C)
FT      old_sequence      /*tag= bj     replace(171, G)
FT      old_sequence      /*tag= bk     replace(174, C)
FT      old_sequence      /*tag= bl     replace(175, C)
FT      old_sequence      /*tag= bm     replace(177, G)
FT      old_sequence      /*tag= bn     replace(180, C)
FT      old_sequence      /*tag= bo     replace(183, G)
FT      old_sequence      /*tag= bp     replace(189, C)
FT      old_sequence      /*tag= bq     replace(192, C)
FT      old_sequence      /*tag= br     replace(198, G)
FT      old_sequence      /*tag= bs     replace(201, G)
FT      old_sequence      /*tag= bt     replace(207, C)
FT      old_sequence      /*tag= bu     replace(210, G)
FT      old_sequence      /*tag= bv     replace(213, C)
FT      old_sequence      /*tag= bw     replace(216, G)
FT      old_sequence      /*tag= bx     replace(219, C)
```


KW Serine protease; ASP05; Insulin-like Growth Factor Binding Protein;
 KW IGFBP; antiproliferative; anti-angiogenic; cancer; ss.
 XX Synthetic.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 CDS 1..1185
 FT /*tag= a
 FT /product= "C-terminal end of serine protease (ASP05) with
 FT Histidine tag"
 XX
 XX W0955885-A2.
 XX
 XX 04-NOV-1999.
 XX
 XX 28-APR-1999; 99WO-US009224.
 XX
 XX 28-APR-1998; 98US-0083321P.
 XX 27-APR-1999; 99US-00300621.
 XX
 XX (AXYS-) AXYS PHARM INC.
 XX
 XX Smeekens SP, Hou J;
 XX
 XX WPI; 2000-072182/06.
 XX P-PSDB; AAY15193.
 XX
 XX New recombinant nucleic acid, with reduced GC content, encoding the
 PT serine protease ASP05, used for treating disorders involving insulin-like
 PT growth factor, e.g. cancer.
 XX
 XX Disclosure; Fig 6; 70pp; English.
 XX
 XX The present sequence is a cDNA encoding the C-terminal domain of serine
 CC protease, ASP05 with a histidine tag. It can be easily purified by
 CC affinity purification using an anti-tag antibody. ASP05 specifically
 CC cleaves insulin-like Growth Factor Binding Protein (IGFBP) and shows
 CC anticancer, antiproliferative and anti-angiogenic activity
 XX
 XX Sequence 1185 BP; 338 A; 291 C; 307 G; 249 T; 0 U; 0 Other;
 SQ
 Query Match 50.8%; Score 1043.8; DB 3; Length 1185;
 Best Local Similarity 95.4%; Pred. No. 5.7e-195;
 Matches 1075; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
 QY 362 CGCAGCGCGGCTGTGTGTGCGCCAGCAGCAGCGCGGTGCGCGCAGCGCAACA 421
 DB 35 CCCAAGCTGGTTATGTGTATGCGCGAGTTCAGAACCAAGTATGTGCTCTGATGCAATA 94
 QY 422 CCTACGCCAACCTGTCAGCTGCGCGCGCGCGCTCCGAGAGCTGCACCGGC 481
 DB 95 CATACGCAAACTTATGCCAATTGAGAGCTGCTTTAGACGTAGTGAAGACTACATAGAC 154
 QY 482 CGCGGTCATGCTCTGCGAGCGGAGCTGCGGCAAGGCGAGGAGGATCCCAACAGTT 541
 DB 155 CGCCTGTATAGTCTTCAACGGGAGGCTGCGGCAAGGCGAGGATCCCAACAGTT 214
 QY 542 TGGCCCAATAATAACTTTATCGGCGAGTGTGAGAGATCGCCCTGCGGTGTTTC 601
 DB 215 TGGCCCAATAATAACTTTATCGGCGAGTGTGAGAGATCGCCCTGCGGTGTTTC 274
 QY 602 ATATCGAATTGTTTCGGAAGCTTCGTTTCTTAAACGAGAGGTGCGGTGAGTGGGT 661
 DB 275 ATATCGAATTGTTTCGGAAGCTTCGTTTCTTAAACGAGAGGTGCGGTGAGTGGGT 334
 QY 662 CTGGGTTTATGTCGGAAGATGACTGATCGTGCACAAATGCCACGTGGTGACCAACA 721
 DB 335 CTGGGTTTATGTCGGAAGATGACTGATCGTGCACAAATGCCACGTGGTGACCAACA 394
 QY 722 AGCACCGGGTCAAAAGTTGAGCTGAAGACGGTGCCACTTACGAAGCCAAATCAAGGATG 781
 DB 395 AGCACCGGGTCAAAAGTTGAGCTGAAGACGGTGCCACTTACGAAGCCAAATCAAGGATG 454

QY 782 TGGATGAGAAAGCAGACATCGCACTCATCAAAATTTGACCACGAGGCAAGCTCCTGTCC 841
 DB 455 TGGATGAGAAAGCAGACATCGCACTCATCAAAATTTGACCACGAGGCAAGCTCCTGTCC 514
 QY 842 TGCTGCTTGGCGGCTCCTCAGAGCTGCGCGGGAGAGTTCTGTGTCGCCATCGGAAGCC 901
 DB 515 TGCTGCTTGGCGGCTCCTCAGAGCTGCGCGGGAGAGTTCTGTGTCGCCATCGGAAGCC 574
 QY 902 CGTTTTCCCTTCAAAACAACAGTCACACCGGGATCGTGAGCACCAACCGAGGCGGCA 961
 DB 575 CGTTTTCCCTTCAAAACAACAGTCACACCGGGATCGTGAGCACCAACCGAGGCGGCA 634
 QY 962 AAGAGCTGGGCTCGGCAACTCAGACATGGACTATCCAGACCGGAGCCATCATCACT 1021
 DB 635 AAGAGCTGGGCTCGGCAACTCAGACATGGACTATCCAGACCGGAGCCATCATCACT 694
 QY 1022 ATGGAACCTCGGAGGCGGCTTAGTAAACCTTGAGCGGTGAAGTGATTGGAATTAACACT 1081
 DB 695 ATGGAACCTCGGAGGCGGCTTAGTAAACCTTGAGCGGTGAAGTGATTGGAATTAACACT 754
 QY 1082 TGAAGTGACAGCTGGATCTCTTTCATCCCTCATCTGATGAAGTATTAAGTTCCTCA 1141
 DB 755 TGAAGTGACAGCTGGATCTCTTTCATCCCTCATCTGATGAAGTATTAAGTTCCTCA 814
 QY 1142 CGGAGTCCCATGACCGACGAGGCAAGGAAAGCCATCACCAAGAAAGTATATTGGTA 1201
 DB 815 CGGAGTCCCATGACCGACGAGGCAAGGAAAGCCATCACCAAGAAAGTATATTGGTA 874
 QY 1202 TCCGAATGATGTCATCTCAGTCCAGCAAGGCAAGAGCTGAAGGACCGGACGGGACT 1261
 DB 875 TCCGAATGATGTCATCTCAGTCCAGCAAGGCAAGAGCTGAAGGACCGGACGGGACT 934
 QY 1262 TCCCAGACGTGATCTCAGGAGCGTATATAATTGAAGTAAATTCCTGATACCCAGCAGAAG 1321
 DB 935 TCCCAGACGTGATCTCAGGAGCGTATATAATTGAAGTAAATTCCTGATACCCAGCAGAAG 994
 QY 1322 CTGGTGTCTTCAAGGAAACGACGTCAATAACAGCATCAATGACAGTGGTGTCTCCG 1381
 DB 995 CTGGTGTCTTCAAGGAAACGACGTCAATAACAGCATCAATGACAGTGGTGTCTCCG 1054
 QY 1382 CCAATGATGTGACGAGCGTCAATTAAGGAAAGCACCCTGAAACATGGTGTCCGAGGG 1441
 DB 1055 CCAATGATGTGACGAGCGTCAATTAAGGAAAGCACCCTGAAACATGGTGTCCGAGGG 1114
 QY 1442 GTAATGAAGATATCATGATCAAGTGTCCGGAAGAAATTCACCA 1488
 DB 1115 GTAATGAAGATATCATGATCAAGTGTCCGGAAGAAATTCACCA 1161

RESULT 13
 ABX77560 standard; cDNA; 1389 BP.
 XX
 XX AC ABX77560;
 XX
 XX DT 09-APR-2003 (first entry)
 XX
 XX DE Differentially expressed breast cancer associated cDNA #55.
 XX
 XX KW Breast cancer; differential gene expression; BC-cDNA;
 KW breast cancer diagnosis; breast cancer monitoring;
 KW breast cancer treatment; breast cancer staging; Gene; ss.
 XX
 XX OS Homo sapiens.
 XX
 XX PN US2002156263-A1.
 XX
 XX PD 24-OCT-2002.
 XX
 XX PF 04-OCT-2001; 2001US-00974298.
 XX
 XX PR 05-OCT-2000; 2000US-0238331P.

XX PA (CHEN/) CHEN H.
 XX PI Chen H;
 XX DR WPI; 2003-182653/18.
 XX PT New cDNAs, which are differentially expressed in (metastatic) breast
 PT cancer useful for diagnosing or staging, breast cancer, or for monitoring
 PT the treatment of breast cancer in an individual.
 XX PS Claim 1; SEQ ID NO 67; 30pp; English.
 XX CC The invention describes a combination of cDNAs (designated BC-cDNAs),
 CC which are differentially expressed in breast cancer. The combination
 CC includes 152 cDNA sequences, or their complements. The protein encoded by
 CC any of these BC-cDNAs is useful for screening several molecules or
 CC compounds to identify at least one ligand that specifically binds the
 CC protein, producing or preparing polyclonal or monoclonal antibodies, or
 CC purifying antibodies from a sample. The antibodies, which specifically
 CC bind the protein differentially expressed in breast cancer is useful for
 CC detecting the expression of a protein in a sample. The BC-cDNAs are also
 CC useful for diagnosing, monitoring the treatment of, or staging, breast
 CC cancer. This sequence represents a differentially expressed breast cancer
 CC associated cDNA. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from the US patent office at
 CC seqdata.uspto.gov/sequence.html?DocID=20020156263
 XX SQ Sequence 1389 BP; 364 A; 324 C; 357 G; 343 T; 0 U; 1 Other;
 Query Match 37.0%; Score 759.8; DB 7; Length 1389;
 Best Local Similarity 99.5%; Pred. No. 2.5e-139;
 Matches 772; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
 QY 1219 ACCTCAGCAAGCAAGAGCTGAAGACCGGACCGGACTTCCACAGCTGATCTCA 1278
 DB 615 ACATCCAGCAAGCAAGAGCTGAAGACCGGACCGGACTTCCACAGCTGATCTCA 674
 QY 1279 GGACGCTATATAATGAAGTAACTCTGATATCCCGACGAGAGCTGGTGTCTCAAGGAA 1338
 DB 675 GGACGCTATATAATGAAGTAACTCTGATATCCCGACGAGAGCTGGTGTCTCAAGGAA 734
 QY 1339 AACGAGCTATATCAGCATCATGAGACAGCTCGGTGTCTCGCCATGATGTCAGGAC 1398
 DB 735 AACGAGCTATATCAGCATCATGAGACAGCTCGGTGTCTCGCCATGATGTCAGGAC 794
 QY 1399 GTCAATTAAGGGAAGCAAGCAAGCTGAGTGGTCCGAGGGGTAAATGAAGATATCATG 1458
 DB 795 GTCAATTAAGGGAAGCAAGCAAGCTGAGTGGTGGTCCGAGGGGTAAATGAAGATATCATG 854
 QY 1459 ATCAGAGTATCCGAGAAATTCACCATAGGACGAGGATGAGCTGGACTTCATGTT 1518
 DB 855 ATCAGAGTATCCGAGAAATTCACCATAGGACGAGGATGAGCTGGACTTCATGTT 914
 QY 1519 TCCCTCAAGAGCTCTCCCGTGATGACGATGAGGACTCTGGGCTCTGCTGAATAGGACAC 1578
 DB 915 TCTCTCAAGAGCTCTCCCGTGATGACGATGAGGACTCTGGGCTCTGCTGAATAGGACAC 974
 QY 1579 TCAGAGCTTTGACGCGCATTTTGTGTTTTCAGTGAGACTCCCTGGCCAGAGATCCT 1638
 DB 975 TCAGAGCTTTGACGCGCATTTTGTGTTTTCAGTGAGACTCCCTGGCCAGAGATCCT 1034
 QY 1639 TCTTGATAGTTGACGCAAAACAAATGTAATGTTTCAGATCCGAGGAGAGAGCTCTGC 1698
 DB 1035 TCTTGATAG-ITGACGCAAAACAAATGTAATGTTTCAGATCCGAGGAGAGAGCTCTGC 1093
 QY 1699 CTTCTGTATCTGATGAGTGTGCTTTTCTTTCAGAGCTTGGCCATCTTCTGTTA 1758
 DB 1094 CTTCTGTATCTGATGAGTGTGCTTTTCTTTCAGAGCTTGGCCATCTTCTGTTA 1153
 QY 1759 GACAGTCAAGCTTTGCTCTCTCTTAAGTATCATCTTATGATCAATCAATGAGT 1818

DB 1154 GACAGTCAGCATTTGTCTCTCTTAACTGAGTCATCATCTTAGTCCCACTAATGAGT 1213
 QY 1819 CGATACAAATCGTAGATAGAGAGAGCCCAACCGGAGCCAGGATGGACTGGTCTGTTTG 1878
 DB 1214 CGATACAAATCGTAGATAGAGAGAGCCCAACCGGAGCCAGGATGGACTGGTCTGTTTG 1273
 QY 1879 TGCCTTTCTCCAAAGTCAGCAGCCCAAGGTCAATGCACAGAGACCCCGGCTGGGTGAGCGC 1938
 DB 1274 TGCCTTTCTCCAAAGTCAGCAGCCCAAGGTCAATGCACAGAGACCCCGGCTGGGTGAGCGC 1333
 QY 1939 TGCCTTTCTCCAAAGTCAGCAGCCCAAGGTGCTCTTTTAGAATCTCTTGAATTTGGAGC 1994
 DB 1334 TGCCTTTCTCCAAAGTCAGCAGCCCAAGGTGCTCTTTTAGAATTTCTTTGAATTTGGAGC 1389
 RESULT 14
 ABQ59513/c
 ID ABQ59513 standard; cDNA; 669 BP.
 AC ABQ59513;
 XX 02-AUG-2002 (first entry)
 DE Human colon cancer related nucleotide sequence SEQ ID NO:3208.
 XX Human; colon cancer; cancer; tissue profiling; forensic; mapping;
 XX Genetic analysis; diagnostic; antisense therapy; gene; ss.
 XX Homo sapiens.
 OS WO200229086-A2.
 FN 11-APR-2002.
 PD 02-OCT-2001; 2001WO-US030732.
 PF 02-OCT-2000; 2000US-0237271P.
 PR (FARB) BAYER CORP.
 PA Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;
 PI Thiagalingam A, Lewis MB;
 PI WPI; 2002-426115/45.
 DR New isolated nucleic acid that is differentially expressed in cancer
 XX tissues useful for determining the presence of colon cancer in a cell or
 PT tissue type, and in antisense therapy.
 PS Claim 1; Fig 1; 796pp; English.
 XX ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
 CC expressed in cancer tissues. ABQ78993 to ABQ79004 represent proteins
 CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
 CC used in antisense therapy. An antibody immunoreactive with a polypeptide
 CC encoded by (I) is useful for detecting a cancer in a patient sample, and
 CC for detecting the presence or absence of a polynucleotide encoded by a
 CC nucleic acid which hybridizes to (I) in a cell. A probe/primer derived
 CC from (I) can be used for determining the presence of a nucleic acid which
 CC hybridizes to (I), and for determining the phenotype of cells in a sample
 CC of cells from a patient. (I) is useful for determining the presence of
 CC colon cancer in a cell or tissue type, for determining the presence or
 CC state of other type of cancer, in antisense therapy, to generate
 CC macroarrays on a solid surface, to identify a chromosome on which the
 CC corresponding gene resides, and in tissue profiling, forensics, genetic
 CC analysis, mapping and diagnostic applications. (I) can be used to raise
 CC antibodies, and to screen for peptide analogues and antagonists
 XX Sequence 669 BP; 177 A; 169 C; 156 G; 167 T; 0 U; 0 Other;
 SQ
 Query Match 30.9%; Score 635; DB 6; Length 669;
 Best Local Similarity 99.6%; Pred. No. 6.1e-115;
 Matches 668; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

Db 760 AATAGAAAAAACCACGGGAACCAAGAGGGACGGGCTTGCTTGAGCCTTACTCCAA 819
 Qy 1892 GTCAGACCCAAAGTCAATGCACAGAGACCCCGGGTGGTGAGCGCTGGCT 1943
 Db 820 GT--AGCCCCCAAGGCTATGCACAGAAACCCCGGTAGGGAAACCTAGACT 869

Search completed: March 10, 2004, 11:00:13
 Job time : 505.903 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 10, 2004, 09:19:42 ; Search time 93.7976 Seconds
(without alignments)
12152.440 Million cell updates/sec

Title: US-10-084-817-32

Perfect score: 2054
Sequence: 1 cgggcctgcctgtccgc.....cattgaaaaaaaaaaaaaa 2054

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:**

- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:**
- 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:**
- 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:**
- 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:**
- 5: /cgn2_6/ptodata/2/ina/PTCUS_COMB.seq:**
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2034	99.0	2036	3	US-08-923-454A-17
2	2010.8	97.9	2205	3	US-08-888-077A-41
3	493.4	24.0	495	4	US-09-702-705-897
4	493.4	24.0	495	4	US-09-736-457-897
5	493.4	24.0	495	4	US-09-614-124B-897
6	493.4	24.0	495	4	US-09-671-325-897
7	479.4	23.3	539	1	US-08-322-742-18
8	419	20.4	441	4	US-09-621-976-14193
9	381.8	18.6	1859	4	US-08-724-864-22
10	317	15.4	358	1	US-08-322-742-19
11	288	14.0	1835	3	US-08-923-454A-7
12	289	14.0	2040	4	US-09-075-460-4
13	288	14.0	2187	3	US-08-923-454A-23
14	288	14.0	2187	3	US-08-923-454A-24
15	288	14.0	2187	3	US-08-923-454A-30
16	288	14.0	2476	3	US-09-008-271A-23
17	224.8	10.9	1503	3	US-08-923-454A-5
18	224.8	10.9	2144	3	US-08-923-454A-28
19	212.4	10.3	1787	3	US-08-923-454A-3
20	212.4	10.3	2551	3	US-08-923-454A-26
21	119	5.8	732	3	US-08-923-454A-1
22	109.8	5.3	1980	1	US-08-350-741-1
23	109.8	5.3	1980	2	US-08-463-875A-1
24	102.4	5.0	1488	4	US-09-489-039A-5255
25	102.4	5.0	1488	4	US-09-489-039A-5193
26	101	4.9	558	4	US-08-976-259-53
27	97	4.7	1854	3	US-08-923-454A-13

28	92.2	4.5	111	1	US-08-322-742-4	Sequence 4, Appl1
29	89.8	4.4	1224	4	US-09-252-931A-8171	Sequence 8171, Ap
30	89.8	4.4	1380	4	US-08-252-931A-7981	Sequence 7981, Ap
31	89.8	4.4	1425	4	US-09-489-039A-6523	Sequence 6523, Ap
32	88.6	4.3	1368	4	US-09-711-164-245	Sequence 245, App
33	88	4.3	1436	4	US-09-199-637A-131	Sequence 131, App
34	87.6	4.3	1083	4	US-09-489-039A-6463	Sequence 6463, Ap
35	84.8	4.1	1428	4	US-09-252-931A-12923	Sequence 12923, A
36	84.8	4.1	1455	4	US-09-252-931A-12561	Sequence 12561, A
37	79.4	3.9	4403765	3	US-09-103-840A-2	Sequence 2, Appl1
38	79.4	3.9	4411529	3	US-09-103-840A-1	Sequence 1, Appl1
39	79.2	3.9	38675	4	US-08-311-731A-135	Sequence 135, App
40	75.6	3.7	891	4	US-09-252-931A-8283	Sequence 8283, Ap
41	71.4	3.5	1771	3	US-08-818-112-13	Sequence 13, Appl
42	71.4	3.5	1771	4	US-08-818-111-13	Sequence 13, Appl
43	71.4	3.5	1771	4	US-09-056-556-13	Sequence 13, Appl
44	71.4	3.5	1771	4	US-09-072-596-13	Sequence 13, Appl
45	71.4	3.5	1771	4	US-09-072-967-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-08-923-454A-17
; Sequence 17, Application US/08923454A
; Patent No. 6004794
; GENERAL INFORMATION:
; APPLICANT: Creasy, Caretha
; APPLICANT: Liivi, George
; APPLICANT: Karan, Eric
; APPLICANT: Clinkbeard, Helen
; APPLICANT: Browne, Michael
; APPLICANT: Southan, Christopher
; TITLE OF INVENTION: HUMAN SERINE PROTEASE
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 1.5
; CURRENT APPLICATION DATA: US/08923,454A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/025436
; FILING DATE: 06-SEPT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Baumeister, Kirk
; REGISTRATION NUMBER: 33,833
; REFERENCE/DOCKET NUMBER: P50547
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5096
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2036 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:

[illegible]

RESULT 2

Db 1456 CGCAGTATGTCAGGAGCTCAATTAAGGAGAAAGCCCTGAAACATGGTGTCCGCAG 1515
 QY 1440 GGGTAATGAAGATATCATGATACAGTGTATCCGAGAAATGACCCATAGCAGAGC 1499
 Db 1516 GGGTAATGAAGATATCATGATACAGTGTATCCGAGAAATGACCCATAGCAGAGC 1575
 QY 1500 ATGAGCTGGACTTCATGTTTCCCTCAAGAGCTCTCCGAGGATGAGGATCTG 1559
 Db 1576 ATGAGCTGGACTTCATGTTTCCCTCAAGAGCTCTCCGAGGATGAGGATCTG 1635
 QY 1560 GGTGCTGGAATAGACACTCAAGACTTTTGTGATGCTGCTATTTTGTGCTGAGGACT 1619
 Db 1636 GGTGCTGGAATAGACACTCAAGACTTTTGTGATGCTGCTATTTTGTGCTGAGGACT 1695
 QY 1620 CCCTGGCCAAAGAGTCTCTTGTGATGTTTGCAGGCAAAACAAATGTAATGTTGCAGAT 1679
 Db 1696 CCCTGGCCAAAGAGTCTCTTGTGATGTTTGCAGGCAAAACAAATGTAATGTTGCAGAT 1755
 QY 1680 CCGCAGGAGAGTCTGCTGCTTCTGATCTCTATGATGAGTGTGCTTTTCTGCCAG 1739
 Db 1756 CCGCAGGAGAGTCTGCTGCTTCTGATCTCTATGATGAGTGTGCTTTTCTGCCAG 1815
 QY 1740 CTGCGGCAATCTGCTAGAGTCAAGTATGCTCTCTTAACTGAGTCAATC 1799
 Db 1816 CTGCGGCAATCTGCTAGAGTCAAGTATGCTCTCTTAACTGAGTCAATC 1875
 QY 1800 TTACTCCAACTAATGAGTCAAGTCAATGCTGATAGAGAGAGCCGAGGAGCAGG 1859
 Db 1876 TTACTCCAACTAATGAGTCAAGTCAATGCTGATAGAGAGAGCCGAGGAGCAGG 1935
 QY 1860 ATGGAGTGGTGTGCTGCTTCTTCTCAAGTCAAGCAGGAGGAGTCAATGAGCAGAG 1919
 Db 1936 ATGGAGTGGTGTGCTGCTTCTTCTCAAGTCAAGCAGGAGGAGTCAATGAGCAGAG 1995
 QY 1920 ACCCGGTGGTGTGCTGCTTCTTCTCAAGCAGGAGGAGTGTGCTTTTGAAGATCTC 1979
 Db 1996 ACCCGGTGGTGTGCTGCTTCTTCTCAAGCAGGAGGAGTGTGCTTTTGAAGATNTN 2055
 QY 1980 TTTGGAATGGGAGCAGATGCTGCTGCTTGTGAGCTATTAAGTACTCTTACACATG 2039
 Db 2056 TTTGGAATGGGAGCAGATGCTGCTGCTTGTGAGCTATTAAGTANTTTTACACATG 2115

RESULT 3
 US-09-702-705-897
 ; Sequence 897, Application US/09702705
 ; Patent No. 6504010
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tonglong
 ; APPLICANT: Bangur, Chaitanya S.
 ; APPLICANT: Lodes, Michael A.
 ; APPLICANT: Fanger, Gary
 ; APPLICANT: Vedvick, Tom
 ; APPLICANT: Carter, Darrick
 ; APPLICANT: Retter, Marc
 ; APPLICANT: Mannion, Jane
 ; APPLICANT: Fan, Liqun
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; FILE REFERENCE: 210121.478C14
 ; CURRENT APPLICATION NUMBER: US/09/702,705
 ; CURRENT FILING DATE: 2000-10-30
 ; NUMBER OF SEQ ID NOS: 1833
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 897
 ; LENGTH: 495
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; US-09-702-705-897

Query Match 24.0%; Score 493.4; DB 4; Length 495;
 Best Local Similarity 99.8%; Pred. No. 5.8e-104;
 Matches 494; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1094 CTGGAATCTCTTTTGAATCCCATCTGATAAGATTAAAAAGTTTCTTCCACGAGTCCCATG 1153
 Db 1 CTGGAATCTCTTTTGAATCCCATCTGATAAGATTAAAAAGTTTCTTCCACGAGTCCCATG 60
 QY 1154 ACCGACAGCCAAAGGAAGCCATCACCAGGAAGATATATTTGGTATCCGATGATGT 1213
 Db 61 ACCGACAGCCAAAGGAAGCCATCACCAGGAAGATATATTTGGTATCCGATGATGT 120
 QY 1214 CACTCACGTCAGCAAAAGCCAAAGAGCTGAAGAGCCGCAACCGGACTTCCACGACGTGA 180
 Db 121 CACTCACGTCAGCAAAAGCCAAAGAGCTGAAGAGCCGCAACCGGACTTCCACGACGTGA 180
 QY 1274 TCTCAGGAGCGTATATTAATTGAGTAATTCCTGATACCCAGGAGCTGGTGGTCTCA 1333
 Db 181 TCTCAGGAGCGTATATTAATTGAGTAATTCCTGATACCCAGGAGCTGGTGGTCTCA 240
 QY 1334 AGGAAACGAGCTCATAATCAGCATCAATGAGCAGTCCGTTCTCGGCAATGATGTCA 1393
 Db 241 AGGAAACGAGCTCATAATCAGCATCAATGAGCAGTCCGTTCTCGGCAATGATGTCA 300
 QY 1394 CGGAGCTATTAAGGGAAGCAGCTGAACATGCTGCTCCGAGGGGTAAGAGATA 1453
 Db 301 CGGAGCTATTAAGGGAAGCAGCTGAACATGCTGCTCCGAGGGGTAAGAGATA 360
 QY 1454 TCATGATCAGAGTGTATCCCGAAGAAATTTGACCCATAGGAGGAGTGTGAGTCTC 420
 Db 361 TCATGATCAGAGTGTATCCCGAAGAAATTTGACCCATAGGAGGAGTGTGAGTCTC 420
 QY 1514 ATGTTTCCCTCAAGACTCTCCGTTGGATGAGGATGAGGACTCTGGGCTGCTGGAATAG 1573
 Db 421 ATGTTTCCCTCAAGACTCTCCGTTGGATGAGGATGAGGACTCTGGGCTGCTGGAATAG 480
 QY 1574 GACACTCAAGACTTTT 1588
 Db 481 GACACTCAAGACTTTT 495

RESULT 4
 US-09-736-457-897
 ; Sequence 897, Application US/09736457
 ; Patent No. 6509448
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tonglong
 ; APPLICANT: Bangur, Chaitanya S.
 ; APPLICANT: Lodes, Michael A.
 ; APPLICANT: Fanger, Gary
 ; APPLICANT: Vedvick, Tom
 ; APPLICANT: Carter, Darrick
 ; APPLICANT: Retter, Marc
 ; APPLICANT: Mannion, Jane
 ; APPLICANT: Fan, Liqun
 ; APPLICANT: Wang, Aijun
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; FILE REFERENCE: 210121.478C15
 ; CURRENT APPLICATION NUMBER: US/09/736,457
 ; CURRENT FILING DATE: 2000-12-13
 ; NUMBER OF SEQ ID NOS: 1864
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 897
 ; LENGTH: 495
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; US-09-736-457-897

Query Match 24.0%; Score 493.4; DB 4; Length 495;
 Best Local Similarity 99.8%; Pred. No. 5.8e-104;
 Matches 494; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1094 CTGGAATCTCTTTTGAATCCCATCTGATAAGATTAAAAAGTTTCTTCCACGAGTCCCATG 1153
 Db 1 CTGGAATCTCTTTTGAATCCCATCTGATAAGATTAAAAAGTTTCTTCCACGAGTCCCATG 60
 QY 1154 ACCGACAGCCAAAGGAAGCCATCACCAGGAAGATATATTTGGTATCCGATGATGT 1213
 Db 61 ACCGACAGCCAAAGGAAGCCATCACCAGGAAGATATATTTGGTATCCGATGATGT 120
 QY 1214 CACTCACGTCAGCAAAAGCCAAAGAGCTGAAGAGCCGCAACCGGACTTCCACGACGTGA 180
 Db 121 CACTCACGTCAGCAAAAGCCAAAGAGCTGAAGAGCCGCAACCGGACTTCCACGACGTGA 180
 QY 1274 TCTCAGGAGCGTATATTAATTGAGTAATTCCTGATACCCAGGAGCTGGTGGTCTCA 1333
 Db 181 TCTCAGGAGCGTATATTAATTGAGTAATTCCTGATACCCAGGAGCTGGTGGTCTCA 240
 QY 1334 AGGAAACGAGCTCATAATCAGCATCAATGAGCAGTCCGTTCTCGGCAATGATGTCA 1393
 Db 241 AGGAAACGAGCTCATAATCAGCATCAATGAGCAGTCCGTTCTCGGCAATGATGTCA 300
 QY 1394 CGGAGCTATTAAGGGAAGCAGCTGAACATGCTGCTCCGAGGGGTAAGAGATA 1453
 Db 301 CGGAGCTATTAAGGGAAGCAGCTGAACATGCTGCTCCGAGGGGTAAGAGATA 360
 QY 1454 TCATGATCAGAGTGTATCCCGAAGAAATTTGACCCATAGGAGGAGTGTGAGTCTC 420
 Db 361 TCATGATCAGAGTGTATCCCGAAGAAATTTGACCCATAGGAGGAGTGTGAGTCTC 420
 QY 1514 ATGTTTCCCTCAAGACTCTCCGTTGGATGAGGATGAGGACTCTGGGCTGCTGGAATAG 1573
 Db 421 ATGTTTCCCTCAAGACTCTCCGTTGGATGAGGATGAGGACTCTGGGCTGCTGGAATAG 480
 QY 1574 GACACTCAAGACTTTT 1588
 Db 481 GACACTCAAGACTTTT 495

QY 1094 CTGGAATCTCTTTGCAATCCCATCTGATAGATTAAAGTTCTCTCAGGAGTCCCATG 1153
Db 1 CTGGAATCTCTTTGCAATCCCATCTGATAGATTAAAGTTCTCTCAGGAGTCCCATG 60
QY 1154 ACCGACAGGCGCAAGGAAAGCCATCCAAAGAAAGATATATTGGTATCCGAATGATG 1213
Db 61 ACCGACAGGCGCAAGGAAAGCCATCCAAAGAAAGATATATTGGTATCCGAATGATG 120
QY 1214 CACTCAGCTCCAGCAAGCCAAAGGCTGAAGGACCGGACCGGACCTTCCAGAGCTGA 1273
Db 121 CACTCAGCTCCAGCAAGCCAAAGGCTGAAGGACCGGACCGGACCTTCCAGAGCTGA 180
QY 1274 TCTCAGGAGCGGTATATTAATTGAAGTAATCTCTGATACCCAGCAAGAGCTGGTCTCA 1333
Db 181 TCTCAGGAGCGGTATATTAATTGAAGTAATCTCTGATACCCAGCAAGAGCTGGTCTCA 240
QY 1334 AGGAAACGAGCTCATATCAGCATCAATGACAGTCCGCTGCTCCGCAATGATGTC 1393
Db 241 AGGAAACGAGCTCATATCAGCATCAATGACAGTCCGCTGCTCCGCAATGATGTC 300
QY 1394 GCGACGTCATTAAAGGAAAGCCCTGAAATCGTGGTCCGAGGGGTAAATGAAGATA 1453
Db 301 GCGACGTCATTAAAGGAAAGCCCTGAAATCGTGGTCCGAGGGGTAAATGAAGATA 360
QY 1454 TCATGATCAGAGTATCCCGAAGAAATTGACCCATAGCAGGACATGAGCTGGACTTC 1513
Db 361 TCATGATCAGAGTATCCCGAAGAAATTGACCCATAGCAGGACATGAGCTGGACTTC 420
QY 1514 ATGTTTCCCTCAAAGACTCTCCCGTGGATGACGATGAGGACTCTGGGCTGTGGAATAG 1573
Db 421 ATGTTTCCCTCAAAGACTCTCCCGTGGATGACGATGAGGACTCTGGGCTGTGGAATAG 480
QY 1574 GACACTCAAGACTTT 1588
Db 481 GACACTCAAGACTTT 495

RESULT 5
US-09-614-124B-897
; Sequence 897, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614.124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 897
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-614-124B-897

Query Match 24.0%; Score 493.4; DB 4; Length 495;
Best Local Similarity 99.8%; Pred. No. 5.8e-104;
Matches 494; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1094 CTGGAATCTCTTTGCAATCCCATCTGATAGATTAAAGTTCTCTCAGGAGTCCCATG 1153
Db 1 CTGGAATCTCTTTGCAATCCCATCTGATAGATTAAAGTTCTCTCAGGAGTCCCATG 60
QY 1154 ACCGACAGGCGCAAGGAAAGCCATCCAAAGAAAGATATATTGGTATCCGAATGATG 1213

Db 61 ACCGACAGGCGCAAGGAAAGCCATCCAAAGAAAGATATATTGGTATCCGAATGATG 120
QY 1214 CACTCAGCTCCAGCAAGCCAAAGGCTGAAGGACCGGACCGGACCTTCCAGAGCTGA 1273
Db 121 CACTCAGCTCCAGCAAGCCAAAGGCTGAAGGACCGGACCGGACCTTCCAGAGCTGA 180
QY 1274 TCTCAGGAGCGGTATATTAATTGAAGTAATCTCTGATACCCAGCAAGAGCTGGTCTCA 1333
Db 181 TCTCAGGAGCGGTATATTAATTGAAGTAATCTCTGATACCCAGCAAGAGCTGGTCTCA 240
QY 1334 AGGAAACGAGCTCATATCAGCATCAATGACAGTCCGCTGCTCCGCAATGATGTC 1393
Db 241 AGGAAACGAGCTCATATCAGCATCAATGACAGTCCGCTGCTCCGCAATGATGTC 300
QY 1394 GCGACGTCATTAAAGGAAAGCCCTGAAATCGTGGTCCGAGGGGTAAATGAAGATA 1453
Db 301 GCGACGTCATTAAAGGAAAGCCCTGAAATCGTGGTCCGAGGGGTAAATGAAGATA 360
QY 1454 TCATGATCAGAGTATCCCGAAGAAATTGACCCATAGCAGGACATGAGCTGGACTTC 1513
Db 361 TCATGATCAGAGTATCCCGAAGAAATTGACCCATAGCAGGACATGAGCTGGACTTC 420
QY 1514 ATGTTTCCCTCAAAGACTCTCCCGTGGATGACGATGAGGACTCTGGGCTGTGGAATAG 1573
Db 421 ATGTTTCCCTCAAAGACTCTCCCGTGGATGACGATGAGGACTCTGGGCTGTGGAATAG 480
QY 1574 GACACTCAAGACTTT 1588
Db 481 GACACTCAAGACTTT 495

RESULT 6
US-09-671-325-897
; Sequence 897, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671.325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 897
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-671-325-897

Query Match 24.0%; Score 493.4; DB 4; Length 495;
Best Local Similarity 99.8%; Pred. No. 5.8e-104;
Matches 494; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1094 CTGGAATCTCTTTGCAATCCCATCTGATAGATTAAAGTTCTCTCAGGAGTCCCATG 1153
Db 1 CTGGAATCTCTTTGCAATCCCATCTGATAGATTAAAGTTCTCTCAGGAGTCCCATG 60
QY 1154 ACCGACAGGCGCAAGGAAAGCCATCCAAAGAAAGATATATTGGTATCCGAATGATG 1213
Db 61 ACCGACAGGCGCAAGGAAAGCCATCCAAAGAAAGATATATTGGTATCCGAATGATG 120
QY 1214 CACTCAGCTCCAGCAAGCCAAAGGCTGAAGGACCGGACCGGACCTTCCAGAGCTGA 1273

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 358
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-322-742-19

Query Match 15.4%; Score 317; DB 1; Length 358;
Best Local Similarity 95.9%; Pred. No. 1.4e-63;
Matches 348; Conservative 0; Mismatches 10; Indels 5; Gaps 2;

QY 1196 TTGGTATCGAATGATGTCACTCAGCTCAGCAAGCCAAAGAGCTGAAGACCGGCACC 1255
DB 1 TTGGTAACCGAATGATGTCACTCAGCTCAGCAAGCCAAAGAGCTGAAGACCGGCACC 60
QY 1256 GGGACTTCCAGACGCTGCTCAGGAGCGTATATAATTGAAGTAATTCCTGATACCCACG 1315
DB 61 GGGACTTCCAGACGCTGCTCAGGAGCGTATATAATTGAAGTAATTCCTGATACCCACG 120
QY 1316 CAGAGCTGGTGGTCTCAGGAAAAACGAGCTCATATATGAGTAAATTCCTGATACCCACG 1375
DB 121 CAGAGCTGGTGGTCTCAGGAAAAACGAGCTCATATATGAGTAAATTCCTGATACCCACG 180
QY 1376 TCTCCGCCAATGATGTCAGGAGCTCATTAAAGGGAAGCACCCTGAACATGGTGTGCC 1435
DB 181 TCTCCGCCAATGATGTCAGGAGCTCATTAAAGGGAAGCACCCTGAACATGGTGTGCC 240
QY 1436 CGAGGGTAATCAAGATATCATGATCACAGTATTCCTCCGGAAGAAATTCACCCATAGGACG 1495
DB 241 GCA-GGGTAATCAAGATATCATGATCACAGTATTCCTCCGGAAGAAATTCACCCATAGGACG 299
QY 1496 AGGCATGAGCTGAGCTTCATGTTCCCTCAAGACTCTCCCGTGGATGACGGATGAGAC 1555
DB 300 AGGCATGAGCTGAGCTTCATGTTCCCTCAAGACTCTCCCGTGGATGACGGATGAGAC 355
QY 1556 TCT 1558
DB 356 TCT 358

RESULT 11

US-08-923-454A-7
Sequence 7, Application US/08923454A
Patent No. 6004794

GENERAL INFORMATION:

APPLICANT: Creasy, Caretha
APPLICANT: Livi, George
APPLICANT: Karan, Eric
APPLICANT: Clinkenbeard, Helen
APPLICANT: Browne, Michael
APPLICANT: Southan, Christopher
TITLE OF INVENTION: HUMAN SERINE PROTEASE
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/923,454A
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/025436
FILING DATE: 06-SEPT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Baumeister, Kirk
REGISTRATION NUMBER: 33,833
REFERENCE/DOCKET NUMBER: P50547
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5096
TELEFAX: 610-270-5090
TELEX:

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:
LENGTH: 1835 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 251...1624
OTHER INFORMATION:
US-08-923-454A-7

Query Match 14.0%; Score 288; DB 3; Length 1835;

Best Local Similarity 57.3%; Pred. No. 1.1e-56;

Matches 541; Conservative 0; Mismatches 400; Indels 3; Gaps 1;

QY 544 CCCCTAAATATAACTTTATTCGCGACGTCGTCGAGAGAGATGCCCTTCGCTGGTTCAT 603
DB 680 CGGAGTCAGTCAACTTTCATCGCAGATGTGTCGAGAGAGACAGCACCCTGCTGCTCTAT 739
QY 604 ATCGAATTTTCGCAAGCTTCGTTTCTAAACGAGAGTCCGCTGGCTAGTGGTCT 663
DB 740 ATCGAGATCCTGGACCGCACCTTTCTTGGCCCGCAGGTCCTATCTCGAACGCTCA 799
QY 664 GGGTTTATTTGTGCGAAGATGAGATGATCGTGACAAATGCCACGTCGTCGTCGTCGTCG 723
DB 800 GGATTCGTGGTGGCTCCGATGGCTCATTTGTACCAACGCCCATGTGGTGGCTGATCGG 859
QY 724 CACCGGTCAAAGTTGAGCTGAAGACGTCGCCATTACGAAGCCAAATCAAGGATGTG 783
DB 860 CGCAGATCCGTGTGAGACTGTCTAAGCGCGACACGATGAGGCGCTGGTCACAGCTGTG 919
QY 784 GATGAGAAAGCAGACATCGCACTCATCAAAATTTGACCACCACGAGGCAAGCTGCTCTCTG 843
DB 920 GATCCCGTGGCAGACATCGCAACGCTGAGGATTCAGACTAAGGAGCCTCTCCCCACGCTG 979
QY 844 CTGCTTGGCGCTCTCTCAGAGCTGGCGCCGGGAGAGTTCGTGTCGCCATCGGAAGCCCG 903
DB 980 CCTCTGGGACGCTCAGCTGATGTCCGGCAAGGGGAGTTTGTGTGCCATGGGAAGTCCC 1039
QY 904 TTTTCCCTTCAAAACACAGTCACACCGGATCGTGAGCACACCCAGCGAGCGGCAAA 963
DB 1040 TTTGCACTGCAGNACAGATCATCGGCAATTTAGTCTCTGTCTCAGGTCACGCCAGA 1099
QY 964 GAGCTGGGCTCGCAACTCAGACATGAGTACATCCAGACCGGACCATCATCACTAT 1023
DB 1100 GACCTGGGACTCCCCCAACCAATGTGGAATACATTCAAACTGATGACGCTATTGATTTT 1159
QY 1024 GGAACTCGGGAGGCGCGTTAGTAAACCTGACCGTGAAGTATTGGGAATTAACACTTTG 1083
DB 1160 GGAACTCTGGAGTCCCTGGTTTACCTGGATGGGAGGTGATTTGGATGAACACCATG 1219
QY 1084 AAAGTCACAGCTGAATCTCTTTTGAATCCATCTCATTAAGATTAAGAGTTCTT---C 1140
DB 1220 AAGGTCACAGCTGAATCTCTTTTGGCATCCCTTCTGATCGTCTTCTCGAGAGTTTCTG 1279

QY 1141 ACGAGTCCATGACCGACAGGCCAAAGGAAAGCCATCACCAGAAAGATATATGGT 1200
DB 1280 CGTGGGAAAGAAATTCCTCTCCGGATCAGTGGTCCAGCGCGTCAATGGG 1339
QY 1201 ATCCGATGATGTCATCAGTCCAGCAAAAGCCAAAGAGCTGAAGAGCCGACCGGAC 1260
DB 1340 GTGATGATGTCAGCCCTGAGTCCAGCATCTTCTGAACTACAGCTTCGAGAACAAGC 1399
QY 1261 TTCCAGAGCTGATCTCAGGAGCGTATATATTAATCAAGTAAATTCCTGATACCCACAGAA 1320
DB 1400 TTCCCGATGTTCCAGATGCTGACTCATCATTAAGTCACTCTGGGCTCCCTGCACAC 1459
QY 1321 GCTGGTGTCTCAAGAAACAGAGTCAATATCAGATCAATATGACATCCGTGCTCTCC 1380
DB 1460 CGGCTGTCTGCGGCTGGTGTGATGATTTTGGCCATTTGGGAGCAGATGGTACAAAAT 1519
QY 1381 GCCAATGATGTGACGACGCTCATTTAAAGGAAAGCAACCTGAACATGGTGGTCCGAGG 1440
DB 1520 GTGAAGATGTTTATGAGCTGTTCGAACCAATCCAGTTGGCAGTGCAGATCCGGCGG 1579
QY 1441 GGTAAATGAAGATATCATGATCAAGTGAATCCCGAAAGAAATGA 1484
DB 1580 GGACGAGAAACACTGACCTTATATGTGACCCCTGAGGTACAGA 1623

RESULT 12

US-09-075-460-4

; Sequence 4, Application US/09075460A

; Patent No. 6489136

; GENERAL INFORMATION:

; APPLICANT: Zervos, Antonis S.

; TITLE OF INVENTION: CELL PROLIFERATION RELATED GENES

; FILE REFERENCE: 10284/004001

; CURRENT APPLICATION NUMBER: US/09/075,460A

; CURRENT FILING DATE: 1998-05-08

; EARLIER FILING DATE: 1997-05-09

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: FASTSEQ for Windows Version 3.0

; SEQ ID NO 4

; LENGTH: 2040

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (248)...(1834)

US-09-075-460-4

Query Match 14.0%; Score 288; DB 4; Length 2040;

Best Local Similarity 57.3%; Pred. No. 1.1e-56;

Matches 541; Conservative 0; Mismatches 400; Indels 3; Gaps 1;

QY 544 CGCCATATAATACTTTATCGGAGCTGGTGGAGAGATCGCCCTCGCGTGGTTCAT 603
DB 890 CGGAGTCAGTACAACTTCATCGAGATGTGGTGGAGAGACAGCACCTCGCGTGGTCTAT 949
QY 604 ATCGATTTGTTCCGAGCTTCGTTTCTTAAAGAGAGGTGCGGCTAGTGGTCT 663
DB 950 ATCGAGATCTCGACCGGACCCCTTCTTGGCGCGGAGGTCCCTATCTCGAACGGCTCA 1009
QY 664 GGGTTTATTTGTGCGAAGATGCACTGATCGTGACAAATGCCACGCTGGTGACCAACAAG 723
DB 1010 GGATTCGTGGTGTGCGGATGGCTCATTTGTACCAACGCCCATGTGGTGGCTGATCGG 1069
QY 724 CACCGGTCAAAGTTGAGCTGAAGACGGTCCCTTACGAGGCAAAATCAAGATGTG 783
DB 1070 CGCAGATCGGTGAGACTGCTAAGCGGGACACGATATGAGGCGGTGGTCAAGCTGTG 1129
QY 784 GATGAGAAAGCAGACATCGCACTCATCAAAATTCACCAAGGCAAGCTGCTGCTCGT 843
DB 1130 GATCCCGTGCAGACATCGCAACGCTGAGGATTCAGACTAAGGACCTCTCCCAAGCTG 1189
QY 844 CTGCTTGGCGCTCTCTCAGAGCTGCGCGGAGAGTTCGTGGTCCCATCGGAAGCCCG 903

DB 1190 CCTCTGGAGCTCAGCTGATGTCCGGCAAGGGAGTTTGTGTGCCATGGAGTCCC 1249
QY 904 TTTTCCCTTCAAACACAGTCCACCGGGATCGGTAGACACCAACCAGGAGCGCGCAAA 963
DB 1250 TTTGCACTGAGAACACGATCACATCCGGCATTTGTAGCTCTGCTCAGCGTCCAGCCAGA 1309
QY 964 GAGCTGGGGTCCGCAACTCAGACATGAGACTACATCCAGACCGACGCCATCATCAACTAT 1023
DB 1310 GACTGGGACTCCCCAAACCAATGTGGAATACATTTCAAACGTGATCGAGCTATTGATTT 1369
QY 1024 GGAACCTGGAGGCCCGTGTAGTAAACCTGACGGTGAAGTGAATTAACACTTTG 1083
DB 1370 GGAACCTGAGAGTCCCTGGTTAAACCTGGATGGGAGGTGATTTGAGTGAACACCATG 1429
QY 1084 AAAGTACAGCTGGAATCTCTTTTGCATCCATCTGATAAGATTAATAAGATTCCT---C 1140
DB 1430 AAGTACAGCTGGAATCTCTTTTGCATCCCTTCTGATCGTCTTCGAGAGTTTCTGCAT 1489
QY 1141 ACGAGTCCCATGACCGACAGGCCAAAGAAAGCCATCAACCAAGAAAGTATATTTG 1200
DB 1490 CGTGGGAAAGAAAGAAATTCCTCTCCGGAATCAGTGGTCCCGCGGCTACATTTGG 1549
QY 1201 ATCCGAATGATGTCACTCAGTCCAGCAAAAGCCAAAGAGCTGAAGACCGGACCGGAC 1260
DB 1550 GTGATGATGTGACCTGAGTCCAGCATCCTTGTCTGACTACAGCTTCGAGAACCAAGC 1609
QY 1261 TTCCAGACGTGATCTCAGGAGCGTATATAATTAAGTAAATTCCTGATACCCAGAGAA 1320
DB 1610 TTTCCGATGTTCCAGCATGTGTGATCTCATCCATAAAGTCACTCTGGGCTCCCTGCACAC 1669
QY 1321 GCTGGTGTCTCAAGAAAGACGAGTCAATATCAGCATCAATGAGACGTCGCTGCTCC 1380
DB 1670 CGGCTGTCTCGGCTGGTGTGATTTTGGCCATTTGGGAGCAGATGTTACAAAT 1729
QY 1381 GCCAATGATCTCAGGAGCTGATTTAAAGGAAAGACCCCTGAAATGTTGTTGTTCCGAGG 1440
DB 1730 GCTGAAGATGTTTATGAAGCTGTTTGAACCAATCCAGTTGGCAGTGCAGATCCGCGG 1789
QY 1441 GGTAAATGAAGATATCATGATCAGATGATTCCTGAGAAATTTGA 1484
DB 1790 GGACGAGAAACACTGACCTTATATGTGACCCCTGAGGTACACAGA 1833

RESULT 13

US-08-923-454A-23

; Sequence 23, Application US/08923454A

; Patent No. 6004794

; GENERAL INFORMATION:

; APPLICANT: Creasy, Caretha

; APPLICANT: Livi, George

; APPLICANT: Karkan, Eric

; APPLICANT: Clinkenbeard, Helen

; APPLICANT: Browne, Michael

; APPLICANT: Southan, Christopher

; TITLE OF INVENTION: HUMAN SERINE PROTEASE

; NUMBER OF SEQUENCES: 40

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Smithline Beecham Corporation

; STREET: 709 Swedeland Road

; CITY: King of Prussia

; STATE: PA

; COUNTRY: USA

; ZIP: 19406

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FASTSEQ Version 1.5

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/923,454A

; FILING DATE:

; CLASSIFICATION: 435

```
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 60/025436
/ FILING DATE: 06-SEPT-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Baumeister, Kirk
/ REGISTRATION NUMBER: 33,833
/ REFERENCE/DOCKET NUMBER: P50547
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 610-270-5096
/ TELEFAX: 610-270-5090
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 23:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2187 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE:
/ ORIGINAL SOURCE:
/ FEATURE:
/ NAME/KEY: Coding Sequence
/ LOCATION: 603...1976
/ OTHER INFORMATION:
/
/ US-08-923-454A-23
/
/ Query Match 14.0%; Score 288; DB 3; Length 2187;
/ Best Local Similarity 57.3%; Pred. No. 1.2e-56;
/ Matches 541; Conservative 0; Mismatches 400; Indels 3; Gaps 1;
/
/ QY 544 GCCATAATAATACTTTATCGGAGCGTGGTGGAGAAGATCGCCCTGCGGTGGTTCAT 603
/ DB 1032 CGGAGTCAGTACAATTCATCGCAGATGGTGGAGAAGACAGCACCTGCGGTGCTAT 1091
/ QY 604 ATCGAATTTTCGAAGCTTCGTTTTCACAGAGAGTGGCGGTGGTGGTCT 663
/ DB 1092 ATCGAGATCTCGGACCGGACCCCTTTCTGGCGCGAGTCCCTATCTCGAAGCGGTCA 1151
/ QY 664 GGGTTTATTGTTCGGAAGATGAGTATCGTGACAAATGCCACGCTGGTGACCAACAAG 723
/ DB 1152 GGAATCGTGGTGGTCCGATGGCTCATGTGTCACCAAGCCCATGTGGTGGTATCGG 1211
/ QY 724 CACCGGTCAAGTTGAGCTGAAGACGGTGCACATACGAAGCCAAATCAGGATGTG 783
/ DB 1212 CGCAGATCCGTGGAGACTGTAAAGCGGACACGTATGAGCGCGTGGTCAAGCTGTG 1271
/ QY 784 GATGAGAAAGCAGACATCGCACTCATCAAAATTGACCAACGAGGCGAAGCTGCTGTCTG 843
/ DB 1272 GATCCCGTGGCAGACATCGCAAGCTGAGATTGAGACTAAGGAGCCTCTCCCAAGCTG 1331
/ QY 844 CTGCTTGGCGCTCTCAGAGCTGGCGCGGAGAGTTCGTGGTCCGATCGGAAGCCG 903
/ DB 1332 CTTCTGGAGCTGAGTCTGATGTCGCGAAGGGAGTTTGTGTCATGGGAGTCCC 1391
/ QY 904 TTTTCCCTTCAAAACACAGTCACCAACCGGATCGTGAGCACCACCCAGCGAGCGGCAAA 963
/ DB 1392 TTTGACATGCAGAACACAGTACATCGGCGATTGTTAGTCTGCTGAGCTCCAGCCAGA 1451
/ QY 964 GAGCTGGGGTCCGCAACTCAGACATGGAATGATCCAGACCGGACCGCATCACTAT 1023
/ DB 1452 GACCTGGGACTCCCCCAACCAATGTGGAAATACATCAAACTGATGAGCTATGATTTT 1511
/ QY 1024 GGAATCTGGGAGGCCGTTAGTAAACCTGGACGTTGAATGATGGAATTAACACTTTG 1083
/ DB 1512 GGAATCTGGAGGTCCCTGTTAACTTGAATGAGGAGTGTGGAGTGAACACCATG 1571
/ QY 1084 AAGGTGAGCTGGATCTCCCTTTGAAATCCCATCTGATAAGATTAAAGATTCCT--C 1140
/ DB 1572 AAGGTGAGCTGGATCTCCCTTTGCAATCCCTTTGATGCTCTGAGAGTTTCTGAT 1631
/ QY 1141 ACGGAGTCCCATGACCGACAGCGCCAAAGGAAAGCCATCAACGAAGAAGATATTGGT 1200
/
/ DB 1632 CGTGGGAAAGAAAGATTCCTCCTCCGGAATCAGTGGGTCCCGCGCGCTACATTGGG 1691
/ QY 1201 ATCCGAATGATGTCTACTCAGTCCAGCAAGCCAAAGAGCTGAAGGACCGGCACGGGAC 1260
/ DB 1692 GTGATGATGCTGACCCCTGAGTCCCGAGCATCCTTTGCTGAATACAGCTTCGAGAACCAGC 1751
/ QY 1261 TTCCAGAGCTGATCTCAGGAGCGTATATAAATTGAAGTAAATTCCTGTATACCCAGCAGAA 1320
/ DB 1752 TTTCCCGATGTTTTCAGCATGGTGTACTCATCCATAAAAGTCAATCCTGGGCTCCCTGCACAC 1811
/ QY 1321 GCTGGTGGTCTCAAGGAAACGACGTCAATATCAGCATCAATGACAGTCCGTGGTCTCC 1380
/ DB 1812 CGGGCTGGTCTCGGCGCTGGTGTGATGATTTTGGCCATTTGGGAGCAGATGGTACAAAT 1871
/ QY 1381 GCCAATGATGTGAGCGACGTCAATTAAGGAAAGCACCCTGAAACATGGTGGTCCGCGAGG 1440
/ DB 1872 GCTGAAGATGTTTATGAAGCTGTTCCGAACCAATCCAGTTGGCAGTGAGATCCGGCGG 1931
/ QY 1441 GGTAAATGAAGATATCATGATCAGATGATTCAGTATCCCGAAGAAATTGA 1484
/ DB 1932 GGAACGAGAAACACTGACCTTATATGTGACCCCTGAGGTACAGA 1975
/
/ RESULT 14
/ US-08-923-454A-24
/ Sequence 24, Application US/08923454A
/ Patent No 6004794
/ GENERAL INFORMATION:
/ APPLICANT: Creasy, Caratha
/ APPLICANT: Livi, George
/ APPLICANT: Karran, Eric
/ APPLICANT: Clinkbeard, Helen
/ APPLICANT: Browne, Michael
/ APPLICANT: Southan, Christopher
/ TITLE OF INVENTION: HUMAN SERINE PROTEASE
/ NUMBER OF SEQUENCES: 40
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: SmithKline Beecham Corporation
/ STREET: 709 Swedeland Road
/ CITY: King of Prussia
/ STATE: PA
/ COUNTRY: USA
/ ZIP: 19406
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq Version 1.5
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/923,454A
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 60/025436
/ FILING DATE: 06-SEPT-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Baumeister, Kirk
/ REGISTRATION NUMBER: 33,833
/ REFERENCE/DOCKET NUMBER: P50547
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 610-270-5096
/ TELEFAX: 610-270-5090
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 24:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2187 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
```

FRAGMENT TYPE:
ORIGINAL SOURCE:
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 603...1976
OTHER INFORMATION:
US-08-923-454A-24

Query Match
Best Local Similarity 14.0%; Score 288; DB 3; Length 2187;
Matches 541; Conservative 0; Mismatches 400; Indels 3; Gaps 1;

QY 544 CGCCATAAATAAATTAACCTTATCGCGGACGCTGGTGAGAGATCGCCCTCGCGTGGTTTCAT 603
DB 1032 CGGAGTCAGTACAACTTCACTCGCAGATGTGGTGAGAGACAGACACCTCGCGTGGTCTAT 1091
QY 604 ATCGAATTTGTTTCGCAAGCTTCGGTTTCTAAACGAGAGGTGCGCGTGGCTAGTGGGTCT 663
DB 1092 ATCGAGATCTCGACCGGACCCCTTCTGGCGCGGAGGTCCCTATCTCGAAGGCTCA 1151
QY 664 GGGTTTATTTGTGCGAAGTGAAGTGAATGATGTCGCAATGCGCAGGTGGTGACCAAG 723
DB 1152 GGATTCGTGGTGGTCCGCGATGGCTCAITGTACCAACGCCCATGTGGTGGCTGATCGG 1211
QY 724 CACCGGGTCAAAAGTTGAGCTGGAAGAACGGTCCCACTTACGAAGCCAAATCAAGGATGTG 783
DB 1212 CGCAGAGTCGGTGGAGCTGCTAAGCGGCGACAGCTATGAGCGCGTGGTTCACAGCTGTG 1271
QY 784 GATGAGAAAGCAGACATCGCACTCAATCAAAATTGACCAACGAGGCAAGCTGCTGCTGTG 843
DB 1272 GATCCGGTGGCAGACATCGCAACGCTGAGGATTCAGACTAAGGAGCCCTCTCCCAACGCTG 1331
QY 844 CTGCTTGGCGCTCCTCAGAGCTGCGGCGCGGAGTTCGTGGTCCCATCGGAAGCCCG 903
DB 1332 CCGTGGGACGCTCAGTGTATGTCGCGCAAGGGAGTTGTTGTGCCATGGAGTCCC 1391
QY 904 TTTTCCCTTCAAAACAGTCACACCGGATCGTGACCAACCCAGCGAGGCGGCAAA 963
DB 1392 TTTGCACTGCAGAACAGTACATCCGCAATGTTAGCTGTCTGTCAGCGTCCAGCCAGA 1451
QY 964 GAGCTGGGGTCCGCAACTCAGACATGAGCTATACATCCAGACCGGCGGCTCATCAACTAT 1023
DB 1452 GACCTGGGACTCCCAACCAATGTGAATACATTCAACTGATGCACTATTGATTTT 1511
QY 1024 GGAAGTGGGAGGCGGTAGTAAACCTCGAGCTGAGTGAATTAACACTTTG 1083
DB 1512 GGAAGTCTGGAGTCCCTGGTTAACTGATGGGAGGTGATGGAGTGAACACCATG 1571
QY 1084 AAAGTGACAGCTGGAATCTCCTTTGCAATCCCATCTGATAGATTAAAGTTCTCT--C 1140
DB 1572 AAGGTACAGCTGGAATCTCCTTTGCCATCCCTCTGATGCTCTGAGAGTTTCTGCAT 1631
QY 1141 ACGGAGTCCCATCGCACAGGCGCAAGGAAAGCCATCACCAGAGAGATATTTGGT 1200
DB 1632 CGTGGGAAAGAAAGAAATTCCTCTCGGAATCAGTGGTCCGAGCGGCTACATTTGG 1691
QY 1201 ATCCGAATGATGTCACTACGTCAGCAAAAGCCAAAGAGCTGAAGGACCGGACCGGAC 1260
DB 1692 GTGATGATGTCAGCCCTGAGTCCAGCATCTTGTGAACTACAGCTTCGAGAACCAAGC 1751
QY 1261 TTCCAGAGCTGATCTCAGGAGCGTATATAATTAAGTAAATTCCTGATACCCGACGAA 1320
DB 1752 TTTCCGAGTTTCAAGATGTTACTCATCAATAAGTCACTCTGGGCTCCCTGCAAC 1811
QY 1321 GCTGGTGGTCTCAAGGAAACGAGCTCATATACGATCAATGAGCAGTCCGTTGCTTCC 1380
DB 1812 CGGGCTGGTCTGCGGCTGGTGTGATTTGTCATTTGGCCATTTGGGAGCAGATGATCAAAAT 1871
QY 1381 GCCAATGATGTGACGAGCGTCATTAAAGGAGAGACCCCTGACATGTTGTTCCGAGG 1440
DB 1872 GCTGAAGATGTTTATGAGCTGTTGCAACCAATCCCAAGTGGGAGTGCAGATCCGCGG 1931
QY 1441 GGTAAATGAAGATATCATGATCAGAGTGAATTCGGAAGAAATTTGA 1484

DB 1932 GGACGAGAAACACTGACCTTATATGTGACCCCTGAGGTACAGAA 1975

RESULT 15
US-08-923-454A-30
Sequence 30, Application US/08923454A
Patent No. 6004794
GENERAL INFORMATION:
APPLICANT: Creasy, Caretha
APPLICANT: Livi, George
APPLICANT: Karian, Eric
APPLICANT: Clinkenbeard, Helen
APPLICANT: Browne, Michael
APPLICANT: Southan, Christopher
TITLE OF INVENTION: HUMAN SERINE PROTEASE
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/923,454A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/025436
FILING DATE: 06-SEPT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Baumeister, Kirk
REGISTRATION NUMBER: 33,833
REFERENCE/DOCKET NUMBER: P50547
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5096
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 2187 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
FEATURE: Polymorphic variants at 672 and 1435
FEATURE: aa24=Arg/cys aa278=Ala/Val
NAME/KEY: Coding Sequence
LOCATION: 603...1976
OTHER INFORMATION:
US-08-923-454A-30

Query Match
Best Local Similarity 14.0%; Score 288; DB 3; Length 2187;
Matches 541; Conservative 0; Mismatches 400; Indels 3; Gaps 1;

QY 544 CGCCATAAATAAATTAACCTTATCGCGGACGCTGGTGAGAGATCGCCCTCGCGTGGTTTCAT 603
DB 1032 CGGAGTCAGTACAACTTCACTCGCAGATGTGGTGAGAGACAGACACCTCGCGTGGTCTAT 1091
QY 604 ATCGAATTTGTTTCGCAAGCTTCGGTTTCTAAACGAGAGGTGCGCGTGGCTAGTGGGTCT 663
DB 1092 ATCGAGATCTCGACCGGACCCCTTCTGGCGCGGAGGTCCCTATCTCGAAGGCTCA 1151

QY	664	GGGTTTATTGTGCGAAGATCGACTGATCGTACAAATGCCACGCTGGTGACCAACAG	723
Db	1152	GGATTGCTGGTGGCTGCCGATGGGCTCATGTGACCAACGCCATGTGGTGGCTGATCGG	1211
QY	724	CACCGGFTCAAAGTTGAGCTGAGAACGGTGGCCACTTACGAAGCCAAATCAAAGGATGTG	783
Db	1212	CGCAGAGTCCGTGTGAGACTGCTAAGCGCGCACGATATGAGGCCGTGGTTCACAGCTGTG	1271
QY	784	GATGAGAAAGCAGACATCGCACTCATCAAAATTTGACCACAGGGCAAGCTGCCTGTCTCG	843
Db	1272	GATCCCGTGGCAGACATGCAACGCTGAGGATTCAGACTAAGAGCCTCTCCCCACGCTG	1331
QY	844	CTGCTTGGCCCGCTCTCAGAGCTGCGCGCGGAGAGTTGCTGGTCCGCATCGGAAGCCCG	903
Db	1332	CCTCTGGACGCTCAGCTGATGTCGGGCAAGGGAGTTTGTGTTGCCATGGGAAGTCCC	1391
QY	904	TTTTCCCTTCAAAACACAGTCAACACCGGGATCGTGAGCACCACCCAGGAGGGCGCAAA	963
Db	1392	TTTGCACTGCGAGAACACGATCACATCCGGCATTGTTAGCTCTCTCAGCGTCCAGCCAGA	1451
QY	964	GAGCTGGGCTCCGCAACTCAGACATGGACTACATCCAGACCGACGCCATCATCACTAT	1023
Db	1452	GACCTGGGACTCCCCCAACCAATGTGGAATACATTCAAACCTGATGACGCTATTGATTTT	1511
QY	1024	GGAAACTCGGAGGCGCCGTTAGTAAACCTGGACGGTGAAGTGAATGGAAATTAACACTTTG	1083
Db	1512	GGAAACTCTGGAGTCCCTGGTTAACTCTGGATGGGAGGTGATTGGAGTGAACACCATG	1571
QY	1084	AAAGTGACAGCTGGAATCTCTTTCCATCCCATCTGATAAGATTAAAGTTTCCT---C	1140
Db	1572	AAGTCAAGTGGAAATCTCTTTGGCATCCCTTCGATCGTCTTCGAGAGTTCTGCAT	1631
QY	1141	ACGAGTCCCATGACCGACAGGCCAAAGAAAGCCATCACCAGAAAGATATATTGGT	1200
Db	1632	CGTGGGAAAGAGAAATTCCTCTCCGGAATCAGTGGGTCCAGCGGGCTACATTGGG	1691
QY	1201	ATCCGAATGATGTCACTCAGCTCCAGCAAGCCAAAGAGCTGAAGACCGGCACCGGAC	1260
Db	1692	GTGATGATGCTGACCTGAGTCCGAGCATCCTTGTGTAATACAGCTTCGAGAACCAAGC	1751
QY	1261	TTCCAGACGTGATCTCAGGAGCGTATATAATTGAAGTAAATTCCTGATACCCGAGAGAA	1320
Db	1752	TTTCCCGATGTTACGATGGTGTACTCATCCATAAAGTCATCCTGGGCTCCCGTGACAC	1811
QY	1321	GCTGGTGTCTCAGGAAACACGCTCATATAATCAGCATCAATGACAGTCCGTTGGTCTCC	1380
Db	1812	CGGCTGTCTCGGCTGGTGTGATGATTTTGGCCATTGGGAGCAGATGGTACAAAT	1871
QY	1381	GCCAAATGATGACGAGCTCATTTAAAGGGAAGCACCCCTGAAATGGTGGTCCGAGG	1440
Db	1872	GCTGAAGATGTTTATGAAGCTGTCGAAACCCCAATCCAGTTGGCAGTGCAGATCCGCGG	1931
QY	1441	GGTAATGAAGATATCATGATCACAGTGAATCCCGAAGAAATGA	1484
Db	1932	GGACGAGAAACACTGACCTTATATGTGACCCCTGAGGTCACAGA	1975

Search completed: March 11, 2004, 14:41:55
Job time : 100.798 secs

Db 1922 TGGAGTGTGTGCTGTTGTGCTTTCTCCAGTCAGACCCAAAGTCAATGACAGAGA 1981
QY 1921 CCGCGGGGGTGAAGCGCTGCTCTCAACGGCGCGAGTTGCTCTTTTGAAGATCTCT 1980
Db 1982 CCGCGGGGGTGAAGCGCTGCTCTCAACGGCGCGAGTTGCTCTTTTGAAGATCTCT 2041
QY 1981 TTGGAATTGGAGACAGATGACTGTGAGTTTGAAGTATTAAAGTACTTCTTACATTTGA 2040
Db 2042 TTGGAATTGGAGACAGATGACTGTGAGTTTGAAGTATTAAAGTACTTCTTACATTTGA 2101
QY 2041 AAAAAAAAAAAAAA 2054
Db 2102 AAAAAAAAAAAAAA 2115

RESULT 3
US-10-264-049-135
; Sequence 135, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birese et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; PRIOR FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 135
; LENGTH: 2157
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1505)..(1505)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2147)..(2147)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-264-049-135

Query Match 99.2%; Score 2037.4; DB 15; Length 2157;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2049; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 1 CCGGCGCTGCGCTGTTCGCGCGCGCAACCGCGCGCGCGAGAGTGCAGATGCCG 60
Db 65 CCGGCGCTGCGCTGTTCGCGCGCGCAACCGCGCGCGCGAGAGTGCAGATGCCG 124
QY 61 CCGGCGCTGTTCGCGCTGTTCGCGCGCGCAACCGCGCGCGCGAGAGTGCAGATGCCG 120
Db 125 CCGGCGCTGTTCGCGCTGTTCGCGCGCGCAACCGCGCGCGCGAGAGTGCAGATGCCG 184
QY 121 TCCGCGCGCGCGCGCTGTTCGCGCGCGCAACCGCGCGCGCGAGAGTGCAGATGCCG 180
Db 185 TCCGCGCGCGCGCGCTGTTCGCGCGCGCAACCGCGCGCGCGAGAGTGCAGATGCCG 244
QY 181 CCGTGTCCGCGCGCGCGAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGC 240
Db 245 CCGTGTCCGCGCGCGCGAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGC 304
QY 241 TGGAGAGTGTGCGCGCGCGCGAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGC 300
Db 305 TGGAGAGTGTGCGCGCGCGCGAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGC 364
QY 301 GAGGAGTGTGAGTGTGCGCGCGCGAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGC 360
Db 365 GAGGAGTGTGAGTGTGCGCGCGCGAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGC 424

QY 361 GCGAGAGCGCGCTGTGTGTGTGCGCGCGAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGC 420
Db 425 GCGAGAGCGCGCTGTGTGTGTGCGCGCGAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGC 483
QY 421 ACCTACGCGCACTGTGTGTGTGCGCGCGAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGC 480
Db 484 ACCTACGCGCACTGTGTGTGTGCGCGCGAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGC 543
QY 481 CCGCGGTGTGTGTGTGTGTGTGCGCGCGAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGC 540
Db 544 CCGCGGTGTGTGTGTGTGTGTGCGCGCGAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGC 603
QY 541 TTGGCGCAATAATAATTAATTTTGT 600
Db 604 TTGGCGCAATAATAATTAATTTTGT 663
QY 601 CATATCGAATTGT 660
Db 664 CATATCGAATTGT 723
QY 661 TGTGGTTTATTTGT 720
Db 724 TGTGGTTTATTTGT 783
QY 721 AAGCAACGGGTCAAGT 780
Db 784 AAGCAACGGGTCAAGT 843
QY 781 GTGATGTGAAGAGAGAGATGT 840
Db 844 GTGATGTGAAGAGAGAGATGT 903
QY 841 CTGCTGT 900
Db 904 CTGCTGT 963
QY 901 CCGTTTTCCTTCAAAACAGATGT 960
Db 964 CCGTTTTCCTTCAAAACAGATGT 1023
QY 961 AAGAGCTGGGCTGT 1020
Db 1024 AAGAGCTGGGCTGT 1083
QY 1021 TATGAAATCTGGAGAGGCGCGTTAGTAAACCTGTGAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
Db 1084 TATGAAATCTGGAGAGGCGCGTTAGTAAACCTGTGAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1143
QY 1081 TTGAAGTGAAGAGT 1140
Db 1144 TTGAAGTGAAGAGT 1203
QY 1141 ACGAGT 1200
Db 1204 ACGAGT 1263
QY 1201 ATCCGATGATGT 1260
Db 1264 ATCCGATGATGT 1323
QY 1261 TTCCCAAGT 1320
Db 1324 TTCCCAAGT 1383
QY 1321 GCTGGT 1380
Db 1384 GCTGGT 1443
QY 1381 GCCAATGATGT 1440
Db 1444 GCCAATGATGT 1503
QY 1441 GGTATGAAGATATCATGTATCATGTATTCGAGAGAAATGATCCATAGGAGAGAGCA 1500

Db	1504	GNTATATGAGATATCATATATACAGCTATATCCCGAAGAAATTGACCCATAGGCAAGGCA	1563
QY	1501	TGAGCTGACATTCATGTTTCCCTCMAAGACTCTCCCGTGATGACGATGAGACTCTGG	1566
Db	1564	TGAGCTGACATTCATGTTTCCCTCAAAAGACTCTCCCGTGATGACGATGAGACTCTGG	1623
QY	1561	GCTGCTGGAAATGAGCACTCAAGACTTTGACTGCAATTTGTTGTTCAATGAGAGATC	1620
Db	1524	GCTGCTGGAAATGAGCACTCAAGACTTTGACTGCAATTTGTTGTTCAATGAGAGATC	1688
QY	1621	CCTGGCCCAAGAACTCTCTTGATAGTTTGACGGCAAAACAATGATATGTCAGATC	1680
Db	1684	CCTGGCCCAAGAAATCCTCTTGATAGTTTGACGGCAAAACAATGATATGTCAGATC	1743
QY	1681	CGAGAGCGAAGAGCTGTGCCCTTCGTATCCATATGATAGACAGTGTCTTTCTTGCCAGC	1746
Db	1744	CGAGAGCGAAGAGCTGTGCCCTTCGTATCCATATGATAGACAGTGTCTTTCTTGCCAGC	1803
QY	1741	TTGGGCCCATTCCTGCTTGAACAGTCAGCAATTTGTCTCCTTTAATCAAGTCATCATCT	1800
Db	1804	TTGGGCCCATTCCTGCTTGAACAGTCAGCAATTTGTCTCCTTTAATCAAGTCATCATCT	1863
QY	1801	TAGTCCAACTATGAGATCGATACAAATCCGATAGATAGAAAGAACCCCAAGGAGCCAGAGA	1863
Db	1864	TAGTCCAACTATGAGATCGATACAAATCCGATAGATAGAAAGAACCCCAAGGAGCCAGAGA	1922
QY	1861	TGGGACTGTGCTGTTTGTGCTTTTCTCCAAATGACGACCCCAAGATCAATGACAGAGA	1922
Db	1924	TGGGACTGTGCTGTTTGTGCTTTTCTCCAAATGACGACCCCAAGATCAATGACAGAGA	1983
QY	1921	CCCCGGGTGGATGAGCGCTGAGCTTCCAAAGCGCGAAGTGGCCCTTTTAGAATCTCT	1988
Db	1984	CCCCGGGTGGATGAGCGCTGAGCTTCCAAAGCGCGAAGTGGCCCTTTTAGAATCTCT	2043
QY	1981	TTGGAAATGGAGCAGATGACTCTGAGTTTGAAGTATTAAGTACTTTTACACATTTGA	2048
Db	2044	TTGGAAATGGAGCAGATGACTCTGAGTTTGAAGTATTAAGTACTTTTACACATTTGA	2103
QY	2041	AAAAAAAAAAAAAA 2054	
Db	2104	AAAAAAAAAAAAAA 2117	

```

RESULT 4
US-09-960-706-516
; Sequence 516, Application US/09960706
; Publication No. US20030134280A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William E.
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplasia
; TITLE OF INVENTION: Gene Expression Profiles
; FILE REFERENCE: 44921-5029-01US
; CURRENT APPLICATION NUMBER: US/09/960,706
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 60/223,323
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 03/873,319
; PRIOR FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 516
; LENGTH: 2036
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134280A1 D87258
US-09-960-706-516

```

Query Match Similarity	99.0%	Score 2034.4	DB 10	Length 2036
Best Local Similarity	100.0%	Pred. No. 0		
Matches 2035	0	Mismatches 1	Indels 0	Gaps 0

[illegible]

QY 841 CTGCTGCTTGGCGGCTCCTCAGAGCTCGGCGCGGAGAGTTCTGGTGCCTATCGAAAGC 900
DB 841 CTGCTGCTTGGCGGCTCCTCAGAGCTCGGCGCGGAGAGTTCTGGTGCCTATCGAAAGC 900
QY 901 CGGTTTTCCTTCAAAACACAGTACCAACCGGATCGTGAGCAACACCCAGGAGCGGC 960
DB 901 CGGTTTTCCTTCAAAACACAGTACCAACCGGATCGTGAGCAACACCCAGGAGCGGC 960
QY 961 AAAGAGCTGGGCTCGCAATCAGACATGAGCTACATCCAGACGAGCGCATCATCAAC 1020
DB 961 AAAGAGCTGGGCTCGCAATCAGACATGAGCTACATCCAGACGAGCGCATCATCAAC 1020
QY 1021 TATGGAACCTCGGAGGCGCGTTAGTAACTGACACGCTGAAGTGAATTAACACT 1080
DB 1021 TATGGAACCTCGGAGGCGCGTTAGTAACTGACACGCTGAAGTGAATTAACACT 1080
QY 1081 TTGAAGTGAACCTGGAATCTCTTTGCAATCCATCTGATAGATTAAGAGTTCTTC 1140
DB 1081 TTGAAGTGAACCTGGAATCTCTTTGCAATCCATCTGATAGATTAAGAGTTCTTC 1140
QY 1141 ACGAGTCCCATCAGCAGAGGCGCAAGGAAAGCCATCACCAGAAAGATATATGGT 1200
DB 1141 ACGAGTCCCATCAGCAGAGGCGCAAGGAAAGCCATCACCAGAAAGATATATGGT 1200
QY 1201 ATCCGAATGATGTCACCTCAGCTCCAGCAAGACGAAAGCTGAAGACCGCACCGGAC 1260
DB 1201 ATCCGAATGATGTCACCTCAGCTCCAGCAAGACGAAAGCTGAAGACCGCACCGGAC 1260
QY 1261 TTCCCAAGCTGATCTCAGAGCGGTATATATTAAGATTAATCTGTATACCCAGAGAA 1320
DB 1261 TTCCCAAGCTGATCTCAGAGCGGTATATATTAAGATTAATCTGTATACCCAGAGAA 1320
QY 1321 GCTGGTGGTCTCAAGGAAACGACGTATATCAGCATCAATGACAGCTCGTGTCTCC 1380
DB 1321 GCTGGTGGTCTCAAGGAAACGACGTATATCAGCATCAATGACAGCTCGTGTCTCC 1380
QY 1381 GCCAATGATCTCAGCAGCTTAAAGGAAAGACCTGACATGCTGTCCGAGG 1440
DB 1381 GCCAATGATCTCAGCAGCTTAAAGGAAAGACCTGACATGCTGTCCGAGG 1440
QY 1441 GGTAAAGAGATATCATGATCAGAGTATCCCGAAGAAATGACCCATAGCAGGCA 1500
DB 1441 GGTAAAGAGATATCATGATCAGAGTATCCCGAAGAAATGACCCATAGCAGGCA 1500
QY 1501 TGAGCTGGACTTCATGTTCCCTCAAGACTCTCCGTTGATGACGGATGAGACTCTGG 1560
DB 1501 TGAGCTGGACTTCATGTTCCCTCAAGACTCTCCGTTGATGACGGATGAGACTCTGG 1560
QY 1561 GCTGCTGGAATAGACACTCAAGACTTTTGTACTGCTATTTTGTTCAGTGGAGCTC 1620
DB 1561 GCTGCTGGAATAGACACTCAAGACTTTTGTACTGCTATTTTGTTCAGTGGAGCTC 1620
QY 1621 CTTGGCCAAACAGAACTCTTGTATGATTTGACGCAAAACAAATGTAATGTCAGATC 1680
DB 1621 CTTGGCCAAACAGAACTCTTGTATGATTTGACGCAAAACAAATGTAATGTCAGATC 1680
QY 1681 CGCAGCAGAAAGCTCTGCCCTCTGTATCTTATGATGATGATGATGATGATGATGATG 1740
DB 1681 CGCAGCAGAAAGCTCTGCCCTCTGTATCTTATGATGATGATGATGATGATGATGATG 1740
QY 1741 TTGGGCCATTTCTTGCTTTAGACAGTCAAGATTTGTCTCTCTTTAACTGAGTCATCT 1800
DB 1741 TTGGGCCATTTCTTGCTTTAGACAGTCAAGATTTGTCTCTCTTTAACTGAGTCATCT 1800
QY 1801 TAGTCCAACTAATGCACTCGATACAACTCGTAGATAGAAAGCCCGCAGGAGCCAGGA 1860
DB 1801 TAGTCCAACTAATGCACTCGATACAACTCGTAGATAGAAAGCCCGCAGGAGCCAGGA 1860
QY 1861 TGGGACTGGTGTGTTTGTGCTTTTCTCCAAATGACGACCCCAAGGTCATGCAAGAGA 1920
DB 1861 TGGGACTGGTGTGTTTGTGCTTTTCTCCAAATGACGACCCCAAGGTCATGCAAGAGA 1920

QY 1921 CCCCGGTGGTGGAGCGCTGCTTCTCAAACGGCGAAGTTGCTCTTTTAGGAATCTCT 1980
DB 1921 CCCCGGTGGTGGAGCGCTGCTTCTCAAACGGCGAAGTTGCTCTTTTAGGAATCTCT 1980
QY 1981 TTGGAATTGGAGCAGCATGACTCTGAGTTTGGAGCTATTAAGTACTTCTTACA 2034
DB 1981 TTGGAATTGGAGCAGCATGACTCTGAGTTTGGAGCTATTAAGTACTTCTTACA 2034

RESULT 6
US-10-102-806-134
; Sequence 134, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 134
; LENGTH: 1855
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1818)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (1845)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-102-806-134

Query Match 86.7%; Score 1780; DB 14; Length 1855;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1802; Conservative 0; Mismatches 0; Indels 2; Gaps 2;
QY 251 GCGGCGCGCGCGAGGCGCGCTGCGGCTTGCAGAGGCGCGTGGCGGCGGCGGCTGC 310
DB 14 GCGGCGCGCGCGAGGCGCGCTGCGGCTTGCAGAGGCGCGTGGCGGCGGCGGCTGC 73
QY 311 AGTGGTGGTCCCTTTCGGGTGCGAGCTCGGCCACGCTGGCGGCGCGCGCGCGCGCG 370
DB 74 AGTGGTGGTCCCTTTCGGGTGCGAGCTCGGCCACGCTGGCGGCGCGCGCGCGCGCG 133
QY 371 GCCTCTGTGTGCGCGCAGCAGCGCGGTGTGCGGAGCGAGCGCAACACCTAGCCCA 430
DB 134 GCCTCTGTGTGCGCCA-CAGCGAGCGCGGTGTGCGGAGCGAGCGCAACACCTAGCCCA 192
QY 431 ACCTGTCCAGCTGCGCGCGCGCGCGCTCGAGAGGCTGACCGCGCGCGCGCGCTCA 490
DB 193 ACCTGTCCAGCTGCGCGCGCGCGCGCTCGAGAGGCTGACCGCGCGCGCGCGCTCA 252
QY 491 TCGTCTGCGAGCGCGAGCTTGGCCCAAGGCGAGAGATCCCAACAGTTTGGCGCAT 550
DB 253 TCGTCTGCGAGCGCGAGCTTGGCCCAAGGCGAGAGATCCCAACAGTTTGGCGCAT 312
QY 551 AATATACTTTATCGCGACGTGTGGAGAGATCGCCCTCGCGTGTTCATATCGAAT 610
DB 313 AATATACTTTATCGCGACGTGTGGAGAGATCGCCCTCGCGTGTTCATATCGAAT 372
QY 611 TGTTCGCAAGCTTCGCTTTTCTAAACGAGAGGTCGCGTGGCTAGTGGGTCTGGGTTTA 670
DB 373 TGTTCGCAAGCTTCGCTTTTCTAAACGAGAGGTCGCGTGGCTAGTGGGTCTGGGTTTA 432
QY 671 TTGTGTCGGAAGATGAGACTGATCTGTGCAAAATGCCCCAAGTGTGACCAACAGCACCGG 730

Db 433 TTGTGTCGGAAGATGAGATGATCGTGAACAAATGCCCACTGGTGACACACACGCGG 492
Qy 731 TCAAAAGTTGAGCTGAAGAACCGTGCCACTTACGAAGCCAAATCAAGGATGTGGATGAGA 790
Db 493 TCAAAAGTTGAGCTGAAGAACCGTGCCACTTACGAAGCCAAATCAAGGATGTGGATGAGA 552
Qy 791 AAGCAGACATCGCATCTCATCAAAATTTGACACAGGCGCAAGCTGCCTGCTGCTGCTG 850
Db 553 AAGCAGACATCGCATCTCATCAAAATTTGACACAGGCGCAAGCTGCCTGCTGCTGCTG 612
Qy 851 GCCGCTCTCAGAGCTGGCGCGGAGAGTTCGTGTCGCCATCGGAAGCCGTTTCC 910
Db 613 GCCGCTCTCAGAGCTGGCGCGGAGAGTTCGTGTCGCCATCGGAAGCCGTTTCC 672
Qy 911 TTCAAAACACAGTACACCGGATCGTGAGACACACAGGCGGCGCAAGAGCTGG 970
Db 673 TTCAAAACACAGTACACCGGATCGTGAGACACACAGGCGGCGCAAGAGCTGG 732
Qy 971 GGCTCCGCAACTCAGACATGGAATATCATCCAGACCGACCATCACTATGGAAC 1030
Db 733 GGCTCCGCAACTCAGACATGGAATATCATCCAGACCGACCATCACTATGGAAC 792
Qy 1031 CGGAGGCGCGTTAGTAACCTGGACGGTGAAGTATGGAATTAACACTTTGAAGTGA 1090
Db 793 CGGAGGCGCGTTAGTAACCTGGACGGTGAAGTATGGAATTAACACTTTGAAGTGA 852
Qy 1091 CAGCTGGAATCTCCTTTGCAATCCCATCTCTATAAGATTAAAAAGTTCTCCACGAGTCCC 1150
Db 853 CAGCTGGAATCTCCTTTGCAATCCCATCTCTATAAGATTAAAAAGTTCTCCACGAGTCCC 912
Qy 1151 ATGACCGACAGCCAAAGGAAAGCCATCACCAGGAAGATATTTGGTATCCGAATGA 1210
Db 913 ATGACCGACAGCCAAAGGAAAGCCATCACCAGGAAGATATTTGGTATCCGAATGA 972
Qy 1211 TGCTACTCAGCTCCAGCAAGCAAGAGCTGAAGACCGGACCGGGAATTTCCACAGC 1270
Db 973 TGCTACTCAGCTCCAGCAAGCAAGAGCTGAAGACCGGACCGGGAATTTCCACAGC 1032
Qy 1271 TGATCTCAGGAGCGTATATAATTGAAGTAACTCTGATACCCAGCAGGAAGCTGGTGGTC 1330
Db 1033 TGATCTCAGGAGCGTATATAATTGAAGTAACTCTGATACCCAGCAGGAAGCTGGTGGTC 1092
Qy 1331 TCAAGGAAACAGCTCATATCAGCATCATATGACAGTCCGTGGTCTCCGCCATGATG 1390
Db 1093 TCAAGGAAACAGCTCATATCAGCATCATATGACAGTCCGTGGTCTCCGCCATGATG 1152
Qy 1391 TCAGCGAGCTCATTAAGGGAAGCACCCTGAACATGGTGGTCCGACGGGTAATGAAG 1450
Db 1153 TCAGCGAGCTCATTAAGGGAAGCACCCTGAACATGGTGGTCCGCA -GGGTAATGAAG 1211
Qy 1451 ATATCATGATCAGTGATTCGAGAGAAATTCACCCATAGGACAGGCAATGAGCTGGAC 1510
Db 1212 ATATCATGATCAGTGATTCGAGAGAAATTCACCCATAGGACAGGCAATGAGCTGGAC 1271
Qy 1511 TTCAATTTCCCTCAAGACTCTCCGCTGATGACCGATGAGACTCTGGCTGCTGGAA 1570
Db 1272 TTCAATTTCCCTCAAGACTCTCCGCTGATGACCGATGAGACTCTGGCTGCTGGAA 1331
Qy 1571 TAGGACACTCAAGACTTTTGACTGCCATTTGTTGTTGAGTGAGACTCCCTGGCCAAAC 1630
Db 1332 TAGGACACTCAAGACTTTTGACTGCCATTTGTTGTTGAGTGAGACTCCCTGGCCAAAC 1391
Qy 1631 AGAATCCTTTCTGATAGTTGAGGCAAAACAAATGTAATTTGACATCGGAGGCGAGA 1690
Db 1392 AGAATCCTTTCTGATAGTTGAGGCAAAACAAATGTAATTTGTCAGATCCGAGGCGAGA 1451
Qy 1691 AGCTCTGCCCTCTGATCTCATGTATGACAGTGTGCTTTTCTTCCAGCAGCTTGGGCCATT 1750
Db 1452 AGCTCTGCCCTCTGATCTCATGTATGACAGTGTGCTTTTCTTCCAGCAGCTTGGGCCATT 1511
Qy 1751 CTTGCTTAGACAGTACAGCTTTGCTCTCTCTTAATGATGATCATCTTAGTCCCACT 1810

Db 1512 CTTGCTTAGACAGTCAGCATTTGTCTCTCTCTTTAACTGAGTCATCATCTTAGTCCAAC 1571
Qy 1811 AATGCACTCGATACAAATGCGTAGATAGAAGAGCCCAAGGAGCCAGGATGGACTGGT 1870
Db 1572 AATGCACTCGATACAAATGCGTAGATAGAAGAGCCCAAGGAGCCAGGATGGACTGGT 1631
Qy 1871 CGTGTCTGTCTTTTCTCAAGTCAGACCCAAAGGTCAATGACAGAGACCCCGGTGG 1930
Db 1632 CGTGTCTGTCTTTTCTCAAGTCAGACCCAAAGGTCAATGACAGAGACCCCGGTGG 1691
Qy 1931 GTGAGCGCTGGCTTCTCAAAACGCGCAAGTTCCTCTTTTAGGAATCTCTTTGGAATTGG 1990
Db 1692 GTGAGCGCTGGCTTCTCAAAACGCGCAAGTTCCTCTTTTAGGAATCTCTTTGGAATTGG 1751
Qy 1991 GAGCAGATGACTCTGAGTTTGAAGTATTAAGTACTCTTACACATTTGAAGAAAAAAA 2050
Db 1752 GAGCAGATGACTCTGAGTTTGAAGTATTAAGTACTCTTACACATTTGAAGAAAAAAA 1811
Qy 2051 AAAA 2054
Db 1812 AAAA 1815

RESULT 7
US-10-104-047-795
; Sequence 795, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1e1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 795
; LENGTH: 1894
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-795

Query Match 82.8%; Score 1701.4; DB 15; Length 1894;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1761; Conservative 0; Mismatches 1; Indels 49; Gaps 1;

Qy 229 GCGTCGCGCTGTCGAGAGTGTGCGCGCCCGAGGCGCGCGTGGCGGCTTGCAGGAG 288
Db 133 GTGTGCGGCTGTCGAGAGTGTGCGCGCCCGAGGCGCGCGTGGCGGCTTGCAGGAG 192
Qy 289 GCGCCGTGCGGCGAGGGGCTGCAAGTGGTGGTCCCTTCGGGGTCCAGCCCTCGGCCACG 348
Db 193 GCGCCGTGCGGCGAGGGGCTGCAAGTGGTGGTCCCTTCGGGGTCCAGCCCTCGGCCACG 252
Qy 349 GTGCGGCGCGCGCGAGCGCGCTTGTGTGTGCGCGAGCAGCGCGGCTGTGCGGC 408
Db 253 GTGCGGCGCGCGCGAGCGCGCTTGTGTGTGCGCGCGCTTGTGTGTGCGCGCAG- 293
Qy 409 AGCGAGCCCAACACCTACGCCACCTGTGTCAGCTGCGCGCCGCGAGCGCGGCTTCCGAG 468
Db 294 -----CAGTCGCGCGCGCGCAGCGCGCTTCCGAG 323
Qy 469 AGGCTGCACCGCGCGCGGTCATCGCTCAGCGCGGAGCTCGCGGCAAGGGCAGGAA 528
Db 324 AGGCTGCACCGCGCGCGGTCATCGCTCAGCGCGGAGGCTCGCGGCAAGGGCAGGAA 383
Qy 529 GATCCCAACAGTTTGGCCATAAATAATTTATTCGCGGACGTTGGTGGAGAAAGTCCGC 588
Db 384 GATCCCAACAGTTTGGCCATAAATAATTTATTCGCGGACGTTGGTGGAGAAAGTCCGC 443
Qy 589 CCGTCGCTGGTTCATATCGAATTTGTTTCGCAAGCTTCCGTTTCTTAACGAGAGTGGCG 648

Db 444 CTTGCGGTGGTTTATATCGAATGTTTTCGCAAGCTTCGGTTTCTAAACGAGAGGTGCGG 503
Qy 649 GTGGCTAGTGGGTCTGGTGTATTTGTGTCGGAAGATGGAATGATCGTGAACAAATGCCAC 708
Db 504 GTGGCTAGTGGGTCTGGTGTATTTGTGTCGGAAGATGGAATGATCGTGAACAAATGCCAC 563
Qy 709 GTGGTGACCAACAAAGCAACCGGGTCAAAAGTTGAGCTGAAGAACGGTGCACATTCAGAAAGCC 768
Db 564 GTGGTGACCAACAAAGCAACCGGGTCAAAAGTTGAGCTGAAGAACGGTGCACATTCAGAAAGCC 623
Qy 769 AAAATCAAGATGTGATGGAAGAGAGATCGCACTCATCAAAATTTGACCAACGAGGC 828
Db 624 AAAATCAAGATGTGATGGAAGAGAGATCGCACTCATCAAAATTTGACCAACGAGGC 683
Qy 829 AAGCTCCCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 888
Db 684 AAGCTCCCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 743
Qy 889 GCATCGGAAGCCGTTTCCCTTCAAAACACAGTCAACACCGGATCGTGAGCAACACC 948
Db 744 GCATCGGAAGCCGTTTCCCTTCAAAACACAGTCAACACCGGATCGTGAGCAACACC 803
Qy 949 CAGCAGGCGGCAAGAGCTGGGGCTCCGCAACTCAGACATGAGCTACATCCAGACCGAC 1008
Db 804 CAGCAGGCGGCAAGAGCTGGGGCTCCGCAACTCAGACATGAGCTACATCCAGACCGAC 863
Qy 1009 GCATCATCAACTATGGAACCTGGGAGCGCGTGTAGTAAACCTGGAGGTGAAGTGATT 1068
Db 864 GCATCATCAACTATGGAACCTGGGAGCGCGTGTAGTAAACCTGGAGGTGAAGTGATT 923
Qy 1069 GGAATTAACACTTTGAAAGTGACAGCTGGAATCTCTTTGCAATCCCATCTGATAAGATT 1128
Db 924 GGAATTAACACTTTGAAAGTGACAGCTGGAATCTCTTTGCAATCCCATCTGATAAGATT 983
Qy 1129 AAAAGTCTCTCAGAGTCCCATGACGAGCGCAGCGCAAGGAAAGCCATCACCAAGAG 1188
Db 984 AAAAGTCTCTCAGAGTCCCATGACGAGCGCAGCGCAAGGAAAGCCATCACCAAGAG 1043
Qy 1189 AAGTATTTGGTATCGAATGATGTCACTCAGCTCCAGCAAGGCAAAAGAGCTGAAGGAC 1248
Db 1044 AAGTATTTGGTATCGAATGATGTCACTCAGCTCCAGCAAGGCAAAAGAGCTGAAGGAC 1103
Qy 1249 CGGACCGGAGCTCCAGAGCTGATCTCAGGAGGTATATAATGAAAGTAACTCTGAT 1308
Db 1104 CGGACCGGAGCTCCAGAGCTGATCTCAGGAGGTATATAATGAAAGTAACTCTGAT 1163
Qy 1309 ACCCCAGCAGAGCTGGTGTCTCAAGGAAACGACGTCATAATCAGCATCAATGACAG 1368
Db 1164 ACCCCAGCAGAGCTGGTGTCTCAAGGAAACGACGTCATAATCAGCATCAATGACAG 1223
Qy 1369 TCGGTGGTCTCCGCAATGATGTCAGGAGCTCATTAAGGGAAGCAACCTGACATG 1428
Db 1224 TCGGTGGTCTCCGCAATGATGTCAGGAGCTCATTAAGGGAAGCAACCTGACATG 1283
Qy 1429 GTGGTCCGAGGGGTAAAGATATCATGATCAGTGATTCCCGAAGAAATGACCCA 1488
Db 1284 GTGGTCCGAGGGGTAAAGATATCATGATCAGTGATTCCCGAAGAAATGACCCA 1343
Qy 1489 TAGGAGAGCATGAGTGAATCATGTTTCCCTCAAGACTCTCCCGTGAATGACGA 1548
Db 1344 TAGGAGAGCATGAGTGAATCATGTTTCCCTCAAGACTCTCCCGTGAATGACGA 1403
Qy 1549 TGAGGACTCTGGGTGCTGGAATAGGACACTCAAGACTTTTGACTGCCATTTGTTGTT 1608
Db 1404 TGAGGACTCTGGGTGCTGGAATAGGACACTCAAGACTTTTGACTGCCATTTGTTGTT 1463
Qy 1609 CAGTGGAGCTCCCTGGCCACAGAACTCTTGTATAGTTTGGAGGCAAAACAAATGTA 1668
Db 1464 CAGTGGAGCTCCCTGGCCACAGAACTCTTGTATAGTTTGGAGGCAAAACAAATGTA 1523
Qy 1669 ATGTTGAGATCCGACGAGAGCTCTGCCCTTCTGTATCTATGATGAGTGTGTT 1728
Db 1524 ATGTTGAGATCCGACGAGAGCTCTGCCCTTCTGTATCTATGATGAGTGTGTT 1583

RESULT 8

US-09-765-231A-20
; Sequence 20, Application US/09765231A
; Patent No. US20020119452A1
; GENERAL INFORMATION:
; APPLICANT: Searle/Monsanto
; APPLICANT: Phippard, Deborah
; APPLICANT: Vasanthakumari, Geetha
; APPLICANT: Dotson, Stanton
; APPLICANT: Ma, Xiao-Jun
; TITLE OF INVENTION: Osteoarthritis tissue-derived nucleic acids, polypeptides,
; TITLE OF INVENTION: vectors, and cells
; FILE REFERENCE: SO-3221 PR
; CURRENT APPLICATION NUMBER: US/09/765,231A
; CURRENT FILING DATE: 2001-01-18
; NUMBER OF SEQ ID NOS: 82
; SEQ ID NO 20
; LENGTH: 1534
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-765-231A-20

Query Match 70.8%; Score 1455.2; DB 9; Length 1534;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1524; Conservative 0; Mismatches 3; Indels 7; Gaps 6;

Qy 525 GGAAGATCCCAACAGTTTGGCCATAAATAAATTAATCGGAGCGTGTGGAGAGAT 584
Db 1 GGAAGATCCCAACAGTTTGGCCATAAATAAATTAATCGGAGCGTGTGGAGAGAT 60
Qy 585 CGCCCTCGCGTGGTTCATATCGAATTTTCGCAAGCTTCGCTTTCTAAACGAGAGGT 644
Db 61 CGCCCTCGCGTGGTTCATATCGAATTTTCGCAAGCTTCGCTTTCTAAACGAGAGGT 120
Qy 645 CGCGTGGCTAGTGGGTCTGGGTTTATGTCGGAAGATGGAATGATCGTGAATGC 704
Db 121 CGCGTGGCTAGTGGGTCTGGGTTTATGTCGGAAGATGGAATGATCGTGAATGC 180
Qy 705 CCACGTGGTGACCAACAAAGCAACCGGGTCAAAGTTGAGCTGAAGAACGGTGCCTTACGA 764
Db 181 CCACGTGGTGACCAACAAAGCAACCGGGTCAAAGTTGAGCTGAAGAACGGTGCCTTACGA 240
Qy 765 AGCCAAATCAAGATGTGATGGAAGCAACATCCACTCATCAAAATGACACCA 824
Db 241 AGCCAAATCAAGATGTGATGGAAGCAACATCCACTCATCAAAATGACACCA 300
Qy 825 GGGCAAGCTGCCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 884

Db	61	TCGGCGCAGCTGTCCTCCGGCGCGCGCTCGCGCGCTTTGGCGCGCGGTGCCAGACCGC	120
Qy	169	TGCGAGCCGGCGCTGCCCGCGCAGCCGAGCACTGCGAGGGCGCGCGCGCGAC	228
Db	121	TGCGAGCCGGCGCTGCGCGCGCAGCCGAGCACTGCGAGGGCGCGCGCGCGCGAC	180
Qy	229	CGCTCGGCTGTCGAGGTGTCGGGCGCGCCGAGGGCGCGCGTGCGGCCTGCAGAG	288
Db	181	CGCTCGGCTGTCGAGGTGTCGGGCGCGCCGAGGGCGCGCGTGCGGCCTGCAGAG	240
Qy	289	GGCCGCTGCGGCGAGGGCTGCAGTGCGTGGTGCCCTTCGGGGTCCAGCCTCGGCCACG	348
Db	241	GGCCGCTGCGGCGAGGGCTGCAGTGCGTGGTGCCCTTCGGGGTCCAGCCTCGGCCACG	300
Qy	349	GTGCGCGCGCGCGCAGGCGCGCTCTGTGTGTGCGCAGCAGCGAGCGCGGTGTGCGGC	408
Db	301	GTGCGCGCGCGCGCAGGCGCGCTCTGTGTGTGCGCAGCAGCGAGCGCGGTGTGCGGC	360
Qy	409	AGCGAGCGCGAACACTTACGCGCAACTGTGCGAGTGTGCGCGCGCGCGCTCCGAG	468
Db	361	AGCGAGCGCGAACACTTACGCGCAACTGTGCGAGTGTGCGCGCGCGCGCTCCGAG	420
Qy	469	AGGCTGACCGCGCGCGCTCATGCTCTGCGAGCGCGAGCGTGTGCGCGCAAGCGCAGGAA	528
Db	421	AGGCTGACCGCGCGCGCTCATGCTCTGCGAGCGCGAGCGTGTGCGCGCAAGCGCAGGAA	480
Qy	529	GATCCCAACAGTTTGGCGCATTAATATAACTTTATCGCGGACGTGGTGGAGAAATCGCC	588
Db	481	GATCCCAACAGTTTGGCGCATTAATATAACTTTATCGCGGACGTGGTGGAGAAATCGCC	540
Qy	589	CCTGCGGTGTTCAATCGAATTTGTTTCGGAAGCTTCGGTTTCTTAAACGAGAGGTGCCG	648
Db	541	CCTGCGGTGTTCAATCGAATTTGTTTCGGAAGCTTCGGTTTCTTAAACGAGAGGTGCCG	600
Qy	649	GTGGTAGTGGGTCTGGGTTTATGTGTGGAAGATGGAATGCTGCGTACAAATGCCAC	708
Db	601	GTGGTAGTGGGTCTGGGTTTATGTGTGGAAGATGGAATGCTGCGTACAAATGCCAC	660
Qy	709	GTGCTGACCAACAAAGCACCGGTCAAAGTTGAGCTGAAGAACGGTGCACATTTACGAAGCC	768
Db	661	GTGCTGACCAACAAAGCACCGGTCAAAGTTGAGCTGAAGAACGGTGCACATTTACGAAGCC	720
Qy	769	AAATCAAGGATGTGATGAGAAGCAGACATCGCACTCATCAAAATTGACCAACAGGCG	828
Db	721	AAATCAAGGATGTGATGAGAAGCAGACATCGCACTCATCAAAATTGACCAACAGGCG	780
Qy	829	AAGCTGCCTGTCTGCTGTGTCGGCTCTCTCAGAGCTGCGGCGCGGAGAGTTCTGTGTC	888
Db	781	AAGCTGCCTGTCTGCTGTGTCGGCTCTCTCAGAGCTGCGGCGCGGAGAGTTCTGTGTC	840
Qy	889	GCCATCGNAGCCGTTTTCCCTTCAAAACACAGTCACCAACGGGATGCTGAGACCAACC	948
Db	841	GCCATCGNAGCCGTTTTCCCTTCAAAACACAGTCACCAACGGGATGCTGAGACCAACC	900
Qy	949	CAGCGAGCGGCAAGAGCTGGGGCTCCGCAACTCAGACATGGAATACATCCAGACCGAC	1008
Db	901	CAGCGAGCGGCAAGAGCTGGGGCTCCGCAACTCAGACATGGAATACATCCAGACCGAC	960
Qy	1009	GCCATCATCAACTATCGAACTCGGGAGCCCGTTAGTAACTTGAACCTGACGGTGAAGTAT	1068
Db	961	GCCATCATCAACTATCGAACTCGGGAGCCCGTTAGTAACTTGAACCTGACGGTGAAGTAT	1020
Qy	1069	GGAATTAACACTTTGAAAGTGACAGCTGGAATCTCTTTGCAATCCCATCTCATAGATT	1128
Db	1021	GGAATTAACACTTTGAAAGTGACAGCTGGAATCTCTTTGCAATCCCATCTCATAGATT	1080
Qy	1129	AAAAAGTTCCTCAGCGAGTCCCATGACCGACAGGCCCAAGGAAGAACATCACCAAGAG	1188
Db	1081	AAAAAGTTCCTCAGCGAGTCCCATGACCGACAGGCCCAAGGAAGAACATCACCAAGAG	1140
Qy	1189	AAGTATATTGGTATCCGAATGATGTCACTCACTGCTCCAGCAAGCCAAAGAGCTGAAGAC	1248
Db	1141	AAGTATATTGGTATCCGAATGATGTCACTCACTGCTCCAGCAAGCCAAAGAGCTGAAGAC	1200

RESULT 10

US-10-198-846-14051
; Sequence 14051, Application US/10198846
; Publication No. US20030099974A1

```

1 / APPLICANT: Lillie, James
2 /
3 / APPLICANT: Xu, Yongyao
4 /
5 / APPLICANT: Wang, Youzhen
6 /
7 / APPLICANT: Steinmann, Kathleen
8 /
9 / TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
10 /
11 / TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
12 /
13 / TITLE OF INVENTION: THERAPY OF BREAST CANCER
14 /
15 / FILE REFERENCE: MRI-049
16 /
17 / CURRENT APPLICATION NUMBER: US/10/199,846
18 /
19 / CURRENT FILING DATE: 2002-07-18
20 /
21 / PRIOR APPLICATION NUMBER: 60/306,220
22 /
23 / PRIOR FILING DATE: 2001-07-18
24 /
25 / NUMBER OF SEQ ID NOS: 14084
26 /
27 / SOFTWARE: FastSEQ for Windows Version 4.0
28 /
29 / SEQ ID NO 14051
30 /
31 / LENGTH: 1498
32 /
33 / TYPE: DNA
34 /
35 / ORGANISM: Homo sapiens
36 /
37 / US-10-198-846-14051

```

Query Match 62.5%; Score 1283.4; DB 14; Length 1498;
Best Local Similarity 91.4%; Pred. No. 0;
Matches 1427. Conservative 0. Mismatches 1. Indels 133.

479	Qy	GGCCGCCGCTCATGCTCTGACGGCGGAGCCTGCGGCCAAGGGCAGGAAGATCCCAACA	538
71	Db	GGCCCGCGTCACTCGTCTGACGGCGGAGCCTGCGGCCAAGGGCAGGAAGATCCCAACA	130
539	Qy	GTTTCGCCATAAATAAATACTTTATCGCGACGCTGCTGGACAAGATCGCCCTGCCGTGG	598
131	Db	GTTTCGCCATAAATAAATACTTTATCGCGACGCTGCTGGACAAGATCGCCCTGCCGTGG	190
599	Qy	TTCATATCGAATTGTTTCGCAAGCTTCCGTTTCTAAACGAGAGGTGCGGTGGCTAGTG	658
191	Db	TTCATATCGAATTGTTTTCGCAAGCTTCCGTTTCTAAACGAGAGGTGCGGTGGCTAGTG	250
659	Qy	GGCTCGGTTTATGTGTCGGAGATGAGTGCATCGTACAAATCCCAACCGTGGTGACCA	718
251	Db	GGCTCGGTTTATGTGTCGGAGATGAGTGCATCGTACAAATCCCAACCGTGGTGACCA	310
719	Qy	ACAAGCACCGGCTCAAAGTTGAGCTGAAGAACCGTGCCACTTACGAAGCAAAATCAAGG	778
311	Db	ACAAGCACCGGCTCAAAGTTGAGCTGAAGAACCGTGCCACTTAC-----	354
779	Qy	ATGTGGATGAGAAAGCAGACATCGCACTCATCAAAATGACCAACCGGCGAAGCTGCCTG	838
355	Db	-----	354
839	Qy	TCCTGCTGTTGGCCGCTCCTCAGAGCTGCGGCCGGGAGAGTTCGTGTCGCGCATCGGAA	898

355 -----GAA 357
 Db 899 GCGCGTTTCCCTTCAAAACACAGTCACACCGGATCGTGACACCAACCGAGGCG 958
 Db 358 GCGCGTTTCCCTTCAAAACACAGTCACACCGGATCGTGACACCAACCGAGGCG 417
 Qy 959 GCAAGAGCTGGGCTCCGCACTCAGACATGAGTACATCAGACCGGACCATCATCA 1018
 Db 418 GCAAGAGCTGGGCTCCGCACTCAGACATGAGTACATCAGACCGGACCATCATCA 477
 Qy 1019 ACTATGAAACTCGGAGCGCGTTAGTAAACCTGGAACCGTGAAGTATGGAAATTAACA 1078
 Db 478 ACTATGAAACTCGGAGCGCGTTAGTAAACCTGGAACCGTGAAGTATGGAAATTAACA 537
 Qy 1079 CTTTGAAGTACAGCTGGAATCTCTTTGCAATCCATCTCATAGATTTAAAGATTTC 1138
 Db 538 CTTTGAAGTACAGCTGGAATCTCTTTGCAATCCATCTCATAGATTTAAAGATTTC 597
 Qy 1139 TCACGGAGTCCCATGACGACGAGCCAAAGGAAAGCCATACCAAGAGAAAGATATATTG 1198
 Db 598 TCACGGAGTCCCATGACGACGAGCCAAAGGAAAGCCATACCAAGAGAAAGATATATTG 657
 Qy 1199 GTATCCGAATCATGTCACTCAGTCCAGCAAGCCAAAGAGCTGAAGGACCGGACCGGG 1258
 Db 658 GTATCCGAATCATGTCACTCAGTCCAGCAAGCCAAAGAGCTGAAGGACCGGACCGGG 717
 Qy 1259 ACTTCCAGAGCTGATCTCAGAGCGGTATATTAATGAAGTAATTCGTATACCCAGCAG 1318
 Db 718 ACTTCCAGAGCTGATCTCAGAGCGGTATATTAATGAAGTAATTCGTATACCCAGCAG 777
 Qy 1319 AAGCTGTGGTCTCAAGGAAACGACGTCATATCAGCATCAATGACACGTCGTCGTCT 1378
 Db 778 AAGCTGTGGTCTCAAGGAAACGACGTCATATCAGCATCAATGACACGTCGTCGTCT 837
 Qy 1379 CCGCAATGATGTGACGACGTCATTTAAAGGAAAGCACCCTGAACATGTTGGTCCGCA 1438
 Db 838 CCGCAATGATGTGACGACGTCATTTAAAGGAAAGCACCCTGAACATGTTGGTCCGCG 897
 Qy 1439 GGGTAATGAGATATCATGATCAGTGATTCGGAAGAAATGACCATAGGACAGG 1498
 Db 898 GGGTAATGAGATATCATGATCAGTGATTCGGAAGAAATGACCATAGGACAGG 957
 Qy 1499 CATGAGCTGGACTCATGTTTCCCTCAAGACTCTCCCGTGGATGCGATGAGGACTCT 1558
 Db 958 CATGAGCTGGACTCATGTTTCCCTCAAGACTCTCCCGTGGATGCGATGAGGACTCT 1017
 Qy 1559 GGGCTGTGGATAGGACACTCAAGACTTTTGAAGTCCCATGTCGATGAGGACTCT 1618
 Db 1018 GGGCTGTGGATAGGACACTCAAGACTTTTGAAGTCCCATGTCGATGAGGACTCT 1077
 Qy 1619 TCCCTGGCCAAACAGAAATCCTTCTGATAGTTTGCAGGCAAAACAAATGTAATGTCAGA 1678
 Db 1078 TCCCTGGCCAAACAGAAATCCTTCTGATAGTTTGCAGGCAAAACAAATGTAATGTCAGA 1137
 Qy 1679 TCCGAGGACAGAGCTCTGCTTCTGATCTCTATGATGACGATGCTTTTCTTGCCA 1738
 Db 1138 TCCGAGGACAGAGCTCTGCTTCTGATCTCTATGATGACGATGCTTTTCTTGCCA 1197
 Qy 1739 GCTTGGGCAATCTTGTCTTAGACAGTCAGCATTTGCTCCTCTCTTAACTGAGTCATCAT 1798
 Db 1198 GCTTGGGCAATCTTGTCTTAGACAGTCAGCATTTGCTCCTCTCTTAACTGAGTCATCAT 1257
 Qy 1799 CTTAGTCCAACTAATGACGTCGATCAATGCGTAGATAGAGAGCCCAACGCGGAGCCAG 1858
 Db 1258 CTTAGTCCAACTAATGACGTCGATCAATGCGTAGATAGAGAGCCCAACGCGGAGCCAG 1317
 Qy 1859 GATGGAGTGTGCTGTTTGTGCTTTTCTCCTAAGTCAGACCCCAAGGTCATGTCACAGA 1918
 Db 1318 GATGGAGTGTGCTGTTTGTGCTTTTCTCCTAAGTCAGACCCCAAGGTCATGTCACAGA 1377
 Qy 1919 GACCCCGGTGGTGGCTGCTGCTTCTCAAGCGGCGGAGTTGCTCTTTTAGGATCT 1978
 Db 1378 GACCCCGGTGGTGGCTGCTGCTTCTCAAGCGGCGGAGTTGCTCTTTTAGGATCT 1437

Qy 1979 CTTTGAATTTGGAGCAGATGACTCTGAGTTTGAAGTATTAAGTACTTCTTACACATT 2038
 Db 1438 CTTTGAATTTGGAGCAGATGACTCTGAGTTTGAAGTATTAAGTACTTCTTACACATT 1497
 Qy 2039 G 2039
 Db 1498 G 1498

RESULT 11
 US-09-974-298-67
 ; Sequence 67, Application US/09974298
 ; Patent No. US20020156263A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Chen, Hui-Mei
 ; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
 ; FILE REFERENCE: PA-0037 P
 ; CURRENT APPLICATION NUMBER: US/09/974,298
 ; CURRENT FILING DATE: 2001-10-04
 ; PRIOR APPLICATION NUMBER: 60/238,331
 ; PRIOR FILING DATE: 2000-05-10
 ; NUMBER OF SEQ ID NOS: 194
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 67
 ; LENGTH: 1389
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No. US20020156263A1 1100429.10
 ; NAME/KEY: unsure
 ; LOCATION: 1372
 ; OTHER INFORMATION: a, t, c, g, or other
 ; US-09-974-298-67

Query Match 37.0%; Score 759.8; DB 9; Length 1389;
 Best Local Similarity 99.5%; Pred. No. 1e-204;
 Matches 772; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 1219 ACGTCCCAAGCAAGCAAGAGCTGAAGGACCGGACCGGACCTCCAGAGCTGATCTCA 1278
 Db 615 ACATCCCAAGCAAGCAAGAGCTGAAGGACCGGACCGGACCTCCAGAGCTGATCTCA 674
 Qy 1279 GGAGCGTATATAATTTGAAGTAATTCCTGATACCCAGCAGAAAGCTGGTCTCAAGGAA 1338
 Db 675 GGAGCGTATATAATTTGAAGTAATTCCTGATACCCAGCAGAAAGCTGGTCTCAAGGAA 734
 Qy 1339 AACGCGTCATAATCAGCATCAATGACAGTCCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1398
 Db 735 AACGCGTCATAATCAGCATCAATGACAGTCCGTCGTCGTCGTCGTCGTCGTCGTCGTC 794
 Qy 1399 GTCATTTAAAGGAAAGCAACCCCTGAACATGTTGGTCCGAGGGGTAAATGAAGATATCATG 1458
 Db 795 GTCATTTAAAGGAAAGCAACCCCTGAACATGTTGGTCCGAGGGGTAAATGAAGATATCATG 854
 Qy 1459 ATCAGAGTATTTCCGAGAAATTTGATGACCATAGGAGGACATGAGTGGACTTCATGTT 1518
 Db 855 ATCAGAGTATTTCCGAGAAATTTGATGACCATAGGAGGACATGAGTGGACTTCATGTT 914
 Qy 1519 TCCCTCAAAGACTCTCCCGTGGATGACGATGAGGACTCTGGGCTGTGGAAATAGGACAC 1578
 Db 915 TCTCTCAAAGACTCTCCCGTGGATGACGATGAGGACTCTGGGCTGTGGAAATAGGACAC 974
 Qy 1579 TCAGAGCTTTTGAAGTGGCTTTTGTGTTTGTTCAGTGGAGACTCCCTGGCCCAACAGATCCT 1638
 Db 975 TCAGAGCTTTTGAAGTGGCTTTTGTGTTTGTTCAGTGGAGACTCCCTGGCCCAACAGATCCT 1034
 Qy 1639 TCTTGATAGTTTGCAGGCAAAACAAATGTAATGTTTCAGATCCGACGAGGAGAGCTTGC 1698
 Db 1035 TCTTGATAG-TTGCAGGCAAAACAAATGTAATGTTTCAGATCCGACGAGGAGAGCTTGC 1093
 Qy 1699 CTTCTGTATCTCTATGATGACAGTGTGCTTTTCTTCCAGCTTGGGCCATTTCTTGCTTA 1758

1094	Db	CTTCTGTATCCATGATGATGAGTGTGCTTTTCTTGGCCAGCTGGGCCATTCTTGCTTA	1153
1759	QY	GACAGTCAGCATTGTCTCTCTCTTTAACTGAGTCATCATCTTAGTCCAACTAAATGCAGT	1818
1154	Db	GACAGTCAGCATTGTCTCTCTCTTTAACTGAGTCATCATCTTAGTCCAACTAAATGCAGT	1213
1819	QY	CGATACAAATCGGTAGATAGAAGAAGCCCAAGCGGAGACGAGTGGGACTCGTCTGTGTTTG	1878
1214	Db	CGATACAAATGGGTAGATAGAAGAAGCCCAAGCGGAGCCAGATGGGACTCGTCTGTGTTTG	1273
1879	QY	TGCTTTTCTCCAGTTCAGACCCCAAGGTCAATGTCACAGAGACCCCGGTGGGTGAGCGC	1938
1274	Db	TGCTTTTCTCCAGTTCAGACCCCAAGGTCAATGTCACAGAGACCCCGGTGGGTGAGCGC	1333
1939	QY	TGGCTTCTCAAAAGCGCCGAAAGTTGCCTCTTTTAGGAAATCTCTTTGGAAATGGGAGC	1994
1334	Db	TGGCTTCTCAAAAGCGCCGAAAGTTGCCTCTTTTAGGAAATCTCTTTGGAAATGGGAGC	1389

```

RESULT 12
US-10-198-846-9582
; Sequence 9582, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MEI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9582
; LENGTH: 797
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2, 3, 6, 9, 17, 630, 633, 706, 728, 784, 789
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-9582

```

Query Match	30.9%	Score 634.4;	DB 14;	Length 797;
Best Local Similarity	96.3%	Pred. NO. 2.9e-169;		
Matches 680;	Conservative	0;	Mismatches 20;	Indels 6;
				Gaps 3;

Qy	1179	CAC	AAG	AAG	AAG	TAT	TGT	GAT	TCC	GAT	TGT	CAC	T	CAC	GT	C	T	AAG	C	AAG	C	AAG	A	1238		
Db	30	CAC	AAG	AAG	AAG	TAT	TGT	GAT	TCC	GAT	TGT	CAC	T	CAC	GT	C	T	AAG	C	AAG	C	AAG	A	89		
Qy	1239	GCT	GAA	GGA	CCG	GAC	CTT	CC	CAG	A	CGT	GAT	TCT	CAG	G	AGC	G	GT	TAT	TAA	T	TGA	AGT	1298		
Db	90	GCT	GAA	GGA	CCG	GAC	CTT	CC	CAG	A	CGT	GAT	TCT	CAG	G	AGC	G	GT	TAT	TAA	T	TGA	AGT	149		
Qy	1299	AAT	T	CCT	GAT	ACC	CC	CAG	CAG	AAG	CTG	TGG	TCT	CAA	G	AAA	AAA	C	AG	C	GT	CAT	TAAT	CAG	AT	1358
Db	150	AAT	T	CCT	GAT	ACC	CC	CAG	CAG	AAG	CTG	TGG	TCT	CAA	G	AAA	AAA	C	AG	C	GT	CAT	TAAT	CAG	AT	209
Qy	1359	CAAT	TGG	A	CAG	T	CCG	CCA	TGA	TGT	CAG	CGA	CGT	CTA	TAA	AAG	G	G	GG	A	AA	AG	CG	AT	1418	
Db	210	CAAT	TGG	A	CAG	T	CCG	CCA	TGA	TGT	CAG	CGA	CGT	CTA	TAA	AAG	G	G	GG	A	AA	AG	CG	AT	269	
Qy	1419	CCT	GAA	CA	TGT	TGT	TCC	CG	AGG	GGT	TA	TGA	AG	A	TAT	CAT	GAT	CA	CAG	T	GA	T	TCC	CG	AA	1478
Db	270	CCT	GAA	CA	TGT	TGT	TCC	CG	AGG	GGT	TA	TGA	AG	A	TAT	CAT	GAT	CA	CAG	T	GA	T	TCC	CG	AA	329

Db 270 CCTGAACATGGTGGTCCGCAGGGGTAATGAAGATATCATGATCACAGTGATTTCCCGAAGA 329

1479	QY	AATTGACCCATAGGCAGAGGCATGAGCTGGACCTTCATGTTTCCCTCAAAGACCTCTCCCGT	1538
330	Db	AAATTGACCCATAGGCAGAGGCATGAGCTGGACCTTCATGTTTCCCTCAAAGACCTCTCCCGT	389
1539	QY	GGATGACGGATCAGGACTCTGGGCTCTGGAAATAGGACACTCAAGACTTTTTGACTGCCAT	1598
390	Db	GGATGACGGATAGGACTCTGGGCTCTGGAAATAGGACACTCAAGACTTTTTGACTGCCAT	449
1599	QY	TTTGTGTTGTCAGTGAGACTCCCTCGCCCAACAGAACTCCTTCGTAGTATTCGACAGCAA	1658
450	Db	TTTGTGTTGTCAGTGAGACTCCCTCGCCCAACAGAACTCCTTCGTAGTATTCGACAGCAA	509
1659	QY	AACAAATGTAATGTTGCAGATCCGACGGCAGAAAGCTCTGCCCTCTGTATTCCTATGTATG	1718
510	Db	AACAAATGTAATGTTGCAGATCCGACGGCAGAAAGCTCTGCCCTCTGTATTCCTATGTATG	569
1719	QY	C-AGTGTCCTTTTTCTTGCCAGC-TTGGGGCCATTCCTGCTTAGACAGTCAGCAATTTGCTCT	1776
570	Db	CAAGTGTCCTTTTTCTTGCCAGCTTTGGGCCATTCCTGCTTAGACAGTCAGCAATTTGCTCT	629
1777	QY	CTTCCTTTAACTAGGTCATCATCTTAGTCGAACTAATGAGTCGATACCAATCGGTAGAT-	1835
630	Db	NCNTCCTTTAACTAGGTCATCATCTTAGTCGAACTAATGAGTCGATACCAATCGGTAGATA	689
1836	QY	---AGAAGAGCCCCACGGAGCCAGGATCGGACTCGTCTGTTTG	1878
690	Db	GGAAAGAGCCCCACCGNAGCCCAAGATGCGGACTCGTNGTGTGTTG	735

RESULT 13
US-10-198-846-9579/c
; Sequence 9579, Application US/10198846
; Publication No. US2003099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongsyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinman, Kathleen
; TITLE OF INVENTION: NOVEL GENES,
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT,
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; TITLE OF INVENTION: PREVENTION, AND

```

FILE REFERENCE: MRI-049
CURRENT APPLICATION NUMBER: US/10/198,846
CURRENT FILING DATE: 2008-07-18
PRIORITY DATE: 2008-07-18
PRIOR FILING DATE: 2008-07-18
NUMBER OF SEQ ID NOS: 14084
SOFTWARE: PastSeq for Windows Version 4.0.
SEQ ID NO 9579
LENGTH: 672

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 14, 21, 23, 27, 29, 30
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-9579

```

Query Match	30.0%	Score 616.8;	DB 14;	Length 672;
Best Local Similarity	99.7%	Pred. No. 2.7e-154;		
Matches 618:	Conservative	0;	Mismatches 2;	Indels 0;
	Gaps	0;		

Qy	1337	AAAAAGCAGCTATAATCAGCATCAATGGACAGTCCGGTGTCTCCGCAATGATGATCGACG	13366
Db	672	AAAAAGCAGCTATAATCAGCATCAATGGACAGTCCGGTGTCTCCGCAATGATGATCGACG	613
Qy	1397	ACGTCAATTAAGGGAAAGCACCCCTGAAACATGTGTGTTCCGACGGGTAAATGAAGATATCA	14566
Db	612	ACGTCAATTAAGGGAAAGCACCCCTGAAACATGTGTGTTCCGACGGGTAAATGAAGATATCA	553
Qy	1457	TGATCACAGTGATTCCCGAAGAAATTGACCCATAGGCAGAGGATGAGCTTGCATTCATG	15166
Db	552	TGATCACAGTGATTCCCGAAGAAATTGACCCATAGGCAGAGGATGAGCTTGCATTCATG	493

1517 TTTCCCTCAAAGACTCTCCCGTGGATGACGAGATGAGGACTCTGGGCTGCTGGGAATAGGAC 1576
492 TTTCCCTCAAAGACTCTCCCGTGGATGACGAGATGAGGACTCTGGGCTGCTGGGAATAGGAC 433
1577 ACTCAAGACTTTTGGACTGCGCAATTTTGGTTTGTTCAGTGGAGACTCCCTGGCCAAACAGATC 1636
432 ACTCAAGACTTTTGGACTGCGCAATTTTGGTTTGTTCAGTGGAGACTCCCTGGCCAAACAGATC 373
1637 CTTCTTGATAGTTTGGAGGCAAAACAAATGTAATGTTTCAGATCCGACGAGGAGAGCTCT 1696
372 CTTCTTGATAGTTTGGAGGCAAAACAAATGTAATGTTTCAGATCCGACGAGGAGAGCTCT 313
1697 GCCCTCTGTATCTTATGATGATGAGTGTCTTTTCTTGGCAGCTTGGCCACTTCTTGCT 1756
312 GCCCTCTGTATCTTATGATGATGAGTGTCTTTTCTTGGCAGCTTGGCCACTTCTTGCT 253
1757 TAGACAGTACGATTTGTCTCTCTCTTTAACTGAGTATCATCTTAGTCCCACTAATGCA 1816
252 TAGACAGTACGATTTGTCTCTCTCTTTAACTGAGTATCATCTTAGTCCCACTAATGCA 193
1817 GTCGATCAATGCTAGATAGAAAGCCCAACGGGAGCCAGGATGGGACTGGTGTGTT 1876
192 GTCGATCAATGCTAGATAGAAAGCCCAACGGGAGCCAGGATGGGACTGGTGTGTT 133
1877 TGTGCTTTTCTCCAAAGTCAGACCCAAAGGTCATGCAAGAGACCCCGGGTGGGTGAGC 1936
132 TGTGCTTTTCTCCAAAGTCAGACCCAAAGGTCATGCAAGAGACCCCGGGTGGGTGAGC 73
1937 GCTGGCTTCTCAAACGGCGC 1956
72 GCTGGCTTCTCAAACGGCGC 53

RESULT 14
US-10-198-846-13646
; Sequence 13646, Application US/10198846
; Publication No. US200309974A1
; GENERAL INFORMATION:
; APPLICANT: Lilie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13646
; LENGTH: 722
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-198-846-13646

Query Match 28.2%; Score 580; DB 14; Length 722;
Best Local Similarity 87.8%; Pred. No. 7.6e-154;
Matches 706; Conservative 0; Mismatches 5; Indels 93; Gaps 3;
1234 AAAGAGCTGAGGACCGGACCGGACTCCAGAGCTGATCTCAGGAGCGTATATAATT 1293
11 AACCGGGGAGGACCGGACCGGACTCCAGAGCTGATCTCAGGAGCGTATATAATT 70
1294 GAAGTAATTCCTGATPACCCAGCAGAGCTGTTGGTCTCAAGGAAACACGCTCATATC 1353
71 GAAGTAATTCCTGATPACCCAGCAGAGCTGTTGGTCTCAAGGAAACACGCTCATATC 130
1354 AGCATCATGACAGCTCGTGGTCTCCGCAATGATGTCAGGACGCTCATTAAGGGAA 1413

131 AGCATCAATGGACAGTCCGTGGTCTCCGCCAATGATGTACGCGAGCTCATTAAGGGAA 190
1414 AGCACCTGAACATGTTGGTCCGAGGGGTAAATGAGATATCATGATCACAGTGATCC 1473
191 AGCACCTGAACATGTTGGTCCGAGGGGTAAATGAGATATCATGATCACAGTGATCC 250
1474 GAAGAAATTCACCCATAGGACGAGCATGAGTGGACTTCATGTTTCCCTCAAAGACTCT 1533
251 GAAGAAATTCACCCATAGGACGAGCATGAGTGGACTTCATGTTTCCCTCAAAGACTCT 310
1534 CCGTGTGATGACGATGAGGACTCTGGGCTGCTGGAATAGGACACTCAAGACTTTGACT 1593
311 CCGTGTGATGACGATGAGGACTCTGGGCTGCTGGAATAGGACACTCAAGACTTTGACT 370
1594 GCCATTTGTTTGTTCAGTGGAGACTCCCTGGCCAAACAGAACTCTTGTATAGTTGCA 1653
371 GCCATTTGTTTGTTCAGTGGAGACTCCCTGGCCAAACAGAACTCTTGTATAGTTGCA 430
1654 GGCNAACAAATGTAATGTTGCAGATCCGAGGAGAGACTCTGCCCTTCTGTATCCTAT 1713
431 GGCNAACAAATGTAATGTTGCAGATCCGCA-GCAGAACTTCTGCTTAGAC 489
1714 GTATGCACTGTGCTTTTCTTCCAGCTTGGGCCAATCTTGTGTTAGACAGTCAGATTG 1773
490 GTATGCACTGTGCTTTTCTTCCAGCTTGGGCCAATCTTGTGCTTAGAC 537
1774 TCTCTCTTAACTGAGTCACTCATCTTAGTCCAACTAATGTCAGTCGATACATGCGTAG 1833
538 ----- 537
1834 ATAGAAGAAAGCCCAAGGAGCCAGGATGGGACTGGTGTGTTGCTTTTCTCCAAGT 1893
538 -----AGCCAGGATGGGACTGGTGTGTTGCTTTTCTCCAAGT 578
1894 CAGC-ACCCAAAGTCAATGACAGAGACCCCGGTTGGGTGAGCGCTGGCTTCTCAAACG 1952
579 CAGCAGCCCAAGGTCATGACAGAGACCCCGGTTGGGTGAGCGCTGGCTTCTCAAACG 638
1953 GCCGAAAGTTGCTCTTTTAGGAATCTCTTTGGAATTTGGAGCACGATGACTCTGAGTTG 2012
639 GCCGAAAGTTGCTCTTTTAGGAATCTCTTTGGAATTTGGAGCACGATGACTCTGAGTTG 698
2013 AGCTATTAAAGTACTTCTTACACA 2036
699 AGCTATTAAAGTACTTCTTACACA 722

RESULT 15
US-10-085-783A-14793
; Sequence 14793, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14793
; LENGTH: 575
; TYPE: DNA
; ORGANISM: Human
US-10-085-783A-14793

Query Match 27.2%; Score 559; DB 12; Length 575;

Best Local Similarity 99.1%; Pred. No. 6.2e-148; Matches 562; Conservative 0; Mismatches 5; Indels 0; Gaps 0;									
Qy	1488	ATAGG	CAGAGG	CATGAG	CTGG	ACTTCAT	GTTCCT	CAAG	ACTCTCCCGTGGATGACGG 1547
Db	1	ATAGG	CAGAGG	CATGAG	CTGG	ACTTCAT	GTTCCT	CAAG	ACTCTCCCGTGGATGACGG 60
Qy	1548	ATGAGG	ACTCTGG	CTGG	CTGG	ATAG	AGAC	CTCA	AGACTTTTGCTGCCATTTGTTTGT 1607
Db	61	ATGAGG	ACTCTGG	CTGG	CTGG	ATAG	AGAC	CTCA	AGACTTTTGCTGCCATTTGTTTGT 120
Qy	1608	TCAGTGG	AGACTCC	CTGG	CAAC	AGAA	TCC	TTTG	ATAGTTTGCAGSCAAAACAAATGT 1667
Db	121	TCAGTGG	AGACTCC	CTGG	CAAC	AGAA	TCC	TTTG	ATAGTTTGCAGSCAAAACAAATGT 180
Qy	1668	AATGTTC	GAGATCCG	CAGG	CAGAG	CTCTG	CCCTT	CTGT	ATCCTATGTATGCAGTGTGCT 1727
Db	181	AATGTTC	GAGATCCG	CAGG	CAGAG	CTCTG	CCCTT	CTGT	ATCCTATGTATGCAGTGTGCT 240
Qy	1728	TTTTCTT	GGCAGCTT	GGGCC	CAATCT	TGCTT	AGAC	AGT	CAGCAATTTGTCTCCTCTTTAAC 1787
Db	241	TTTTCTT	GGCAGCTT	GGGCC	CAATCT	TGCTT	AGAC	AGT	CAGCAATTTGTCTCCTCTTTAAC 300
Qy	1788	TGAGTCA	TCACTT	AGTCC	AACTA	ATC	AGT	CGAT	CAATCGCTAGATAGAGAGCCCC 1847
Db	301	TGAGTCA	TCACTT	AGTCC	AACTA	ATC	AGT	CGAT	CAATCGCTAGATAGAGAGCCCC 360
Qy	1848	ACGGGAG	CCAGG	ATGG	AGCTGG	TGCTG	TTTGT	GTCTT	TCTCAAGTCAGCACCCCAAGGT 1907
Db	361	ACGGGAG	CCAGG	ATGG	AGCTGG	TGCTG	TTTGT	GTCTT	TCTCAAGTCAGCACCCCAAGGT 420
Qy	1908	CAATGC	CACAG	AGACCC	CGGGT	GGGTG	AGCG	CTTCT	CNAACGGCCGAGTTGCCCTCT 1967
Db	421	CAATGC	CACAG	AGACCC	CGGGT	GGGTG	AGCG	CTTCT	CNAACGGCCGAGTTGCCCTCT 480
Qy	1968	TTTAGGA	ATCTCT	TTTGA	ATTGG	AGC	AGAT	GACT	CTGAGTTTTCAGCTATTAAGTACT 2027
Db	481	TTTAGGA	ATCTCT	TTTGA	ATTGG	AGC	AGAT	GACT	CTGAGTTTTCAGCTATTAAGTACT 540
Qy	2028	TCTTACA	CATTTG	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA 2054
Db	541	TCTTACA	CATTTG	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA 567

Search completed: March 12, 2004, 06:15:07
Job time : 484.733 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 10, 2004, 09:18:38 ; Search time 3400.62 Seconds
(without alignments)
18036.975 Million cell updates/sec

Title: US-10-084-817-32

Perfect score: 2054

Sequence: 1 cgggcctgcctgcgcg.....cattgaaaaaaaaaaaaa 2054

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- EST:*
- 1: em_estba:*
 - 2: em_esthum:*
 - 3: em_estin:*
 - 4: em_estnu:*
 - 5: em_estov:*
 - 6: em_estpl:*
 - 7: em_estro:*
 - 8: em_htc:*
 - 9: gb_esti:*
 - 10: gb_est2:*
 - 11: gb_htc:*
 - 12: gb_est3:*
 - 13: gb_est4:*
 - 14: gb_est5:*
 - 15: em_estfun:*
 - 16: em_estom:*
 - 17: em_gss_hum:*
 - 18: em_gss_inv:*
 - 19: em_gss_pln:*
 - 20: em_gss_vrt:*
 - 21: em_gss_fun:*
 - 22: em_gss_man:*
 - 23: em_gss_mus:*
 - 24: em_gss_pro:*
 - 25: em_gss_rod:*
 - 26: em_gss_phg:*
 - 27: em_gss_vrl:*
 - 28: gb_gss1:*
 - 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1703.6	82.9	1911	11 BC011352	BC011352 Homo sapi
2	1173	57.1	1753	11 AK090320	AK090320 Mus muscu
C 3	1051.2	51.2	1201	13 BX334954	BX334954 BX334954
C 4	1048.8	51.1	1201	13 BX356240	BX356240 BX356240

RESULT 1
BC011352
LOCUS
DEFINITION
IMAGE:4177882), containing frame-shift errors.
ACCESSION
BC011352.1 GI:15030191
KEYWORDS
HTC.
SOURCE
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1911)
AUTHORS
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uedin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Pailey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,

ALIGNMENTS

5	1027.4	50.0	1201	13	BX356241	BX356241
6	1021.6	49.7	1146	13	BX423030	BX423030
C 7	1016.8	49.5	1132	13	BX423029	BX423029
C 8	1013.6	49.3	1201	13	BX381058	BX381058
C 9	994	48.4	1090	13	BX423083	BX423083
C 10	988.8	48.1	1201	13	BX339704	BX339704
C 11	973.8	47.4	1106	13	BX423084	BX423084
C 12	973	47.4	973	29	AY406047	AY406047
C 13	941.4	45.8	992	13	BX401342	BX401342
C 14	940.2	45.8	1201	13	BX334955	BX334955
C 15	938.4	45.7	1201	13	BX380623	BX380623
C 16	935.4	45.5	1201	9	AL533136	AL533136
C 17	929.6	45.3	1178	13	BX381846	BX381846
C 18	919	44.7	1201	13	BX417479	BX417479
C 19	915.4	44.6	1001	9	AL568473	AL568473
C 20	914	44.5	1201	13	BX439802	BX439802
C 21	912	44.4	1201	13	BX396059	BX396059
C 22	907.6	44.2	1007	9	AL540553	AL540553
C 23	901.4	43.9	1201	13	BX379690	BX379690
C 24	896.6	43.7	940	13	BX340279	BX340279
C 25	891.4	43.4	1201	9	AL540253	AL540253
C 26	889.4	43.3	1147	13	BX381845	BX381845
C 27	882.4	43.0	1201	9	AL571001	AL571001
C 28	875	42.6	1201	13	BX339705	BX339705
C 29	869	42.3	1147	14	CK232729	CK232729
C 30	867	42.2	1201	9	AL540254	AL540254
C 31	866.8	42.2	1006	13	BX419376	BX419376
C 32	864.8	42.1	1201	9	AL564978	AL564978
C 33	864.4	42.1	950	9	AL542818	AL542818
C 34	863.2	42.0	1201	13	BX381059	BX381059
C 35	860.8	41.9	1201	13	BX334775	BX334775
C 36	853.6	41.6	913	9	AL569639	AL569639
C 37	850.6	41.4	1147	9	AL550616	AL550616
C 38	840.2	40.9	902	9	AL578595	AL578595
C 39	839	40.8	1026	13	BX401343	BX401343
C 40	839	40.8	1201	13	BX334774	BX334774
C 41	834.2	40.6	919	9	AL556304	AL556304
C 42	832	40.5	974	9	AL566107	AL566107
C 43	831.6	40.5	1201	13	BX364740	BX364740
C 44	829.2	40.4	1201	13	BX343677	BX343677
C 45	829	40.4	881	14	CDS18132	CDS18132

BC011352 1911 bp mRNA linear HTC 19-NOV-2003
Homo sapiens protease, serine, 11 (IGF binding), mRNA (cDNA clone
IMAGE:4177882), containing frame-shift errors.

JOURNAL REFERENCE	Nature 409, 685-690 (2001)
AUTHORS	The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
REFERENCE	6 (bases 1 to 1753)
AUTHORS	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T., Horii,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
TITLE	Direct Submission
JOURNAL	Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL:http://genome.gsc.riken.go.jp/ URL:http://fantom.gsc.riken.go.jp/ Location/Qualifiers 1. .1753 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM DB:G630051C16" /db_xref="MGI:2429226" /db_xref="taxon:10090" /clone="G630051C16" /sex="female" /tissue_type="ovary and uterus" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="10 days pregnant adult"
FEATURES	<1..1227 /notes="unnamed protein product; protease, serine, 11 (Igfbinding) (MGD MGI:1929076, GB NM_019564, evidence: BLASTN, 98%, match=1749) putative" /codon_start=1 /protein_id="BAC41168.1" /db_xref="GI:26355508" /translation="GAACGLQGPGCGEGLCQTVFPVPASATVRRRAQLCVCASEPVGCSDATYTNLQCLRAARSRKLPVFIVLQAGCGQEDNSLRHKNFIADVEKTPAVTHIELTKLPFSREVFVSGLVEDGLIIVTNNKVKKVFLKNGATEAKIKVDKERADIALKDHKGLPVLGRSSBELREFGVATGPSFSLQNTTGVTGSTQRGGELGRNSMDIQTDALINYGSGPLVNLDEVIIGTLTKVTAAGISPAIDPKIKFTSKARKGKAVTKKYIGIRMMSLTSXAKELDKHRDPFDVLSGAYITEVPTDPAEAGGLKENDVIISINSQSVVTANDVSDVIKKENTLNVRFRGNEDITVTPBEIDP"
CDS	polyA_signal 1753 polyA_site 1753 note="putative" note="putative"
ORIGIN	Query Match 57.1%; Score 1173; DB 11; Length 1753; Best Local Similarity 82.8%; Pred. No. 1.2e-179;

Query Match 57.1%; Score 1173; DB 11; Length 1753;
Best Local Similarity 82.8%; Pred. No. 1.2e-179;

	Matches 1467;	Conservative	0;	Mismatches	275;	Indels	29;	Gaps	10;
QY	265	GGCGCGCGTGC	GGCGCGTGCAGAGAGGCGCGTGCAGAGAGGCGTGCAGTGCAGTGCAGTGC	CC	324				
DB	1	GTGTGACAGCTG	CGCGCTGCAGAGAGGTCCTCGCGGCGAGGGGTGCAATGCTAGTGC	CC	60				
QY	325	TTCCGGGTGCCAGCCT	CGGCGCACGCTGCGCGCGCGCGCGCAGCGCCGCGCTCTCTGTGTGC	384					
DB	61	TTCCGGGTGCGGCGCT	CGGCCACAGTACGACGCGCGCGCACAGCGCGCGCTGTGCGTGTGT	120					
QY	385	GGCAGCAGGAGCGCGGTG	GGCAGCAGGAGCGCAACACTACGCCAACCTGTCGCGAGCTGTCGCGAGCTG	444					
DB	121	GCCAGCAGGAGCGCGTGT	GGTGTAGCGAGCGCCAGACCTTACACCAACCTGTGCCAGCTG	180					
QY	445	CGCGCGCGCAGCGCGCT	CCGAGAGGCTGCACCGCGCGCGGTCTGCTCTCGAGCGC	504					
DB	181	CGCGCGCGCAGCGCGCT	CCGAGAGGCTTCCGAGCGCGCGGTCTGCTCTCGAGCGC	240					
QY	505	GGAGCTCGCGCCAAAGGG	CAGAGAGTCCCAACAGTTTGGCGCATAAATATACTTTATC	564					
DB	241	GGCGCTCGCGCCAAAGG	CAGAGTCCCAACAGTTTGGCTATAAGTACAACTTTATT	300					
QY	565	GGGACGTGTGTGAGAG	AGATCGCCCGCTCGCGTTCATACGAATGTTTCCGAGCTT	624					
DB	301	GCTGATGTGTGTGAG	AGATCGCCCGCTGTGGTTCACATTGAACATATATCCGAGCTT	360					
QY	625	CGCTTTTCTAAACGAG	AGGTCCGCGTGGCTAGTGGGTCTGGGTCTTATTTGTGTCCGAGAT	684					
DB	361	CGTTTCTGAAGAGGAG	GTGCCAGTGGCGAGTGGGTTCAGGATTCATCGTATCCGAGGAT	420					
QY	685	GGACTGATCGTGACAA	TGCCACGCTGGTGACCAACAAAGCACCGGTCAAAGTTTGAGCTG	744					
DB	421	GGACTGATTTGTGAC	AAATGCTACGTTGGTGACCAACAAACACCGGCTCAAGTTTGAGCTG	480					
QY	745	AAGAACGTTGCCACTTA	CGAAGCCAAATCAAGGATGTGGATGAGAAAGCAGACATCGCA	804					
DB	481	AAGAATTGGAGCTAC	TATGTAAGCCAAATCAAGGATGTGGATGAAAGGCGGACATTCGC	540					
QY	805	CTCATCAAAATTGAC	ACCACAGGCGAAGCTCCCTGTCTGCTGTGTGGCGCCCTCTCTCAGAG	864					
DB	541	CTTATCAAGATTAC	CACAAAGGAAAGCTCCAGTCTCTGCTGTCTGGCGCGCTCTCTCAGAG	600					
QY	865	CTGCGCGCGGAGTTC	GTGTTGTCCTATCGGAGCGCGTTTCCCTTCAAAACACAGTCTC	924					
DB	601	CTGAGACCTGGAGAA	TTGTGTGTTGCCATTTGGAAGCCCTTTCTCTTCAAAACACAGTCTC	660					
QY	925	ACCACCGGATCTGAC	ACCAACCGAGGCGGCAAGAGCTGGGGCTCCGCAACTCA	984					
DB	661	ACCACCTGGATCTG	ACGACCAACCGAGGCGGCAAGAGCTGGGACCTTCGGAATCC	720					
QY	985	GACATGACATACAT	CAGACCGGCGCATCATCAACTATGAAACTCGGAGAGCCCGTTA	1044					
DB	721	GATATGGACTACAT	TCCAGACAGCGCTATCATCAATTTATGAAATTCGAGAGCCCGTTA	780					
QY	1045	GTAAACCTGACCGGT	GAAGTATGGAAATTAACACTTTGAAAGTGACAGCTGGAAATCTCC	1104					
DB	781	GTAAACCTGGATG	CGGAGGTGATTTGGATTTAAACCTTTGAAGGTGACGGCGGGCATCTCC	840					
QY	1105	TTTTCGAATCCCAT	CTCTGATAAGATTAATAAGTTCTCTCACGGAGTCCCATGACCGACAGGCC	1164					
DB	841	TTTCGAATTCAT	CCGATAAGATAAAAGATTTCTTGACAGAGTCCCAACGATCGACAGGCC	900					
QY	1165	AAAGGAAAGCCCAT	CACCAAGAGAAGTATATTGGTATCCGAATGATGTCACTTCACGTCTC	1224					
DB	901	AAAGGNAAGCTGT	CACCAAGAGAAGTATATTGGGATCCGAATGATGTCTCATCATCT	960					
QY	1225	AGCAAGCCAAAGCT	GAAGGACCGGACCGGGAATTTCCAGACGTGATCTCAGGAGCG	1284					
DB	961	AGCAAGCCAAAGCT	GAAGGACCGGTCACCGAGCTTCCCGGATGTCTCTCTGGGGCA	1020					
QY	1285	TATATATTGAAGT	TAATCTCTGTATACCCCAAGCAAGCTGGTGTCTCAAGGAAACGAC	1344					
DB	1021	TATATCATTTGA	GTCAATCTCTGACACCCCGGACAGCCGAGGGCTCAAGGAAATGAC	1080					

1345	Qy	GTGATTAATGAGCATCAATGAGACAGTCTGTTGGTCTCCGCCAATGATGTGACGGAGCTCAATT	1400
1081	Db	GTGATCATGAGCATCAACGGACAGTCTGTGGTCACTGCCAATGACGTGACGGATGTGCATC	1140
1405	Qy	AAAAGGGAAGCACCCCTGAAACATGGTGGTCCGAGGGGTAAATGAAGATATCATGATCACAA	1464
1141	Db	AAAAGGAGAAACACCCTGAAACATGGTGTGCCGAGGGGCAATGAAGACATTTGTGATTACC	1200
1465	Qy	GTGATTTCCCGAAGAAATGACCCATAGGCAGAGGCGATGAGCTGAGCTTCATGTTTCCCTC	1524
1201	Db	GTGATTTCTGAAGAAATCGACCCCTAGGCAGAGGACGAGACCTTCATGTTTCCCTC	1260
1525	Qy	AAAGACTCTCCCGTGGATGACGAGATGAGGACTCTGGCTCTGTGGAATAGGACACTCAAGA	1584
1261	Db	AAAGA-CTCCAGGGATGGCGATGAGAA-CTGAGACTGGTGCACGAGACACCCGGGA	1316
1585	Qy	CTTTTGACTGCCATTTGTTGTTTCAAGTGAGAGACTCCCTGGCCAAACAGAAATCCTTCTTGA	1644
1317	Db	CTTTTGTCACAACTTTTGCTTGTTCA--GGGAACACCCCTGCCGACAGGATCCTTCTTGA	1374
1645	Qy	TAGTTTGACGGCAAAACAAATGTAATGTTTCAGATCCGAGGCGAGAGCTCTGCCCTTCT	1704
1375	Db	TAGTTTGCGGGCAAAACAAATGTAATGTTTCAGAGTGCACAGGACAGAAATCTG-CCTTCT	1433
1705	Qy	GTATCTCATGTATGCAAGTGTGCTTTTTCTTGGCAGCTTGGGCCATCTTCTGTTAGACAGT	1764
1434	Db	GTATGCTATGTATGACAGCTGCTTTTTCCTTACAAGCTTGGGCTGTTCTGCTGTACACAGT	1493
1765	Qy	CAGATTGTGCTCCTCTCCTTTAACTGAGTCAATCTTTAGTCCAACTAATGCAGTCGATAC	1824
1494	Db	CAACATTTGCTCCTTCCCTTAGC-----CTAAGTGTCTGACTAATGCAGTTGACGG	1545
1825	Qy	AATGCGTAGATAGAGAAGCCCCACGGGAGCCAGGATGGGACTGGTCTGTTTGTGCTTT	1884
1546	Db	ATGCGTAGGCAAGAGGAAGTCTCTC-----CGAGCCATGGTCTTGGGCGTGTTTACACTTT	1601
1885	Qy	T-CTCAAGTCAGCACCCCAAAG--TCAATGCACAGACACCCCGGCTGGGTGAGCGCTG	1940
1602	Db	TCCTCCGAGTCGGCACTCAGAGGAATTTGATGCCCGAGACCAAGGTGGGTGA-CCCTG	1650
1941	Qy	GCCTCTCAACGCGCGAAGTGGCTCTTTAGGAATCTCTTTGGAAATTTGGAGACCAAGT	2000
1661	Db	GCCTCTGAGATGCCAAAATGCTCTT-----AGGAATCCTGGAACCCGGAGTACAGTG	1715
2001	Qy	ACTCTGAGTTTGAGCTATTAAAGTACTTCTT	2031
1716	Db	ACTTTGAGTTTCAGCTATTAAATACTCTT	1746

RESULT 3
BX334954/c
LOCUS
DEFINITION BX334954 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CSODI009YE13 3-PRIME, mRNA sequence.
ACCESSION BX334954
VERSION BX334954.1 GI:30343366
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5631.f. For
more information about this cluster, see

digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN	
Query Match	51.1%; Score 1048.8; DB 13; Length 1201;
Best Local Similarity	93.3%; Pred. No. 1.4e-159;
Matches 1087; Conservative	27; Mismatches 48; Indels 3; Gaps 3;
QY	841 CTGCTGCTTGGCGCTCTCTCAGAGCTGGCGCGGAGAGTTCGTCGTCCTCAGAGC 900
DB	1163 CTGCTGCTTGGCGCTCTCTCAGAGCTGGCGCGGAGAGTTCGTCGTCCTCAGAGC 1105
QY	901 CCCTTTTCCCTTCAAAAACACAGTACACACCGGATCGTGAGCACCACCCAGGCGCGC 960
DB	1104 CCCTTTTCCCTTCAAAAACACAGTACACACCGGATCGTGAGCACCACCCAGGCGCGC 1045
QY	961 AAAGAGCTGGGGCTCCGCACTCAGACATCGACTACATCCAGACGAGCCCATCATCAAC 1020
DB	1044 AAAGAGCTGGGGCTCCGCACTCAGACATCGACTACATCCAGACGAGCCCATCATCAAC 985
QY	1021 TATGGAACCTCGGAGGCGCTTGTAGTAAACCTGGAAGTGAAGTGAAGTGAAGTGAAGT 1080
DB	984 TATGGAACCTCGGAGGCGCTTGTAGTAAACCTGGAAGTGAAGTGAAGTGAAGTGAAGT 926
QY	1081 TTGAAAGTGCAGCTGGAAATCTCTTTGCAATCCCATCTGATAAGATTAAAGATTTCCTC 1140
DB	925 TTGAAAGTGCAGCTGGAAATCTCTTTGCAATCCCATCTGATAAGATTAAAGATTTCCTC 866
QY	1141 ACGAGTCCCATGACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200
DB	865 ACGAGTCCCATGACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 806
QY	1201 ATCCGAATGATGTCACTCAGCTCCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1260
DB	805 ATCCGAATGATGTCACTCAGCTCCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 746
QY	1261 TTCCAGACGTCATCTAGGAGGATATATTAAGTGAATTCCTGATATCCCTGATATCCCTG 1320
DB	745 TTCCAGACGTCATCTAGGAGGATATATTAAGTGAATTCCTGATATCCCTGATATCCCTG 686
QY	1321 GCTGGTGTCTCAAGGAAACAGCTCATATCAGCATCAATGGACGAGTCGCTGCTCC 1380
DB	685 GCTGGTGTATCAAGGAAACAGCTCATATCAGCATCAATGGACGAGTCGCTGCTCC 626
QY	1381 GCCAAT-GATCTCAGCGAGCTCATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1439
DB	625 ACCAATAGATGTCCGAGCAACWAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 566
QY	1440 GGGAATGAAGATATCATGATCAGATGATTCGGAAGAAATGACCCATAGGAGGAGGC 1499
DB	565 GGGAATGAAGATATCATGATCAGATGATTCGGAAGAAATGACCCATAGGAGGAGGC 506
QY	1500 ATGAGCTGGACTTCATGTTTCCCTCAAGACTCTCCGCGGATGAGGAGGAGGAGGAGGAG 1559
DB	505 ATGAGCTGGACWACATGTTTCCCAACAGACTCTCCGCGGATGAGGAGGAGGAGGAGGAG 446
QY	1560 GGCTGTGGAATAGGACACTAAGACTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1619
DB	445 GGCAGCTGGATAGGACACTAAGACTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 386
QY	1620 CCCTGGCCAAACAGATCCCTTCTGATGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1679
DB	385 CCCTGGCCAAACAGATCCCTTCTGATGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 326
QY	1680 CCGAG 1739
DB	325 CCGAG 266
QY	1740 CTTGGGCCATTTCTGTTAGACAGTCAAGATTTGCTCCTCCTTTAACTGAGTCATCATC 1799
DB	265 CTTGGGCAATTTCTGTTAKACAGTCAAGATTTGCTCCTCCTTTAACTGAGTCATCATC 206
QY	1800 TTAGTCCCACTAATGCACTGATCAATGCGTAGATAGAGAGGAGGAGGAGGAGGAGGAGGAG 1859

DB	205 TTAGTCCCACTAATGCACTGATCAATCGTAGATAGAGAGCCCAACCGGAGCAGG 146
QY	1860 ATGGAGTGTGCTGTTTGTCTTTTCTCCCAAGTCAGCACCACCAAGGTCAATGCACAGAG 1919
DB	145 ATGGAGTGTGCTGTTTGTCTTTTCTCCCAAGTCAGCACCACCAAGGTCAATGCACAGAG 86
QY	1920 ACCCGCGGTGGTGAGCGCTGCTTCTCAAAACGCGGAGGAGTTCCTTTTAGGAATCTC 1979
DB	85 ACCCGCGGTGGTGAGCGCTGCTTCTCAAAACGCGGAGGAGTTCCTTTTAGGAATCTC 26
QY	1980 TTTGGAATTTGGAGCAGCATGACTC 2004
DB	25 TNNGAATTTGGAGCAGCATGACTC 1
RESULT 5	
LOCUS	BX356241 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION	Clone CSODI009Y110 5-PRIME, mRNA sequence.
ACCESSION	BX356241
VERSION	BX356241.1 GI:30382060
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	1 (bases 1 to 1201)
AUTHORS	Li W.B., Gruber C., Jessee J. and Polayes D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seq@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 5631.f For more information about this cluster, see http://www.genoscope.cns.fr/ cgi-bin/cluster.cgi?seq=CSODI009DF05QP1&cluster=5631.f. Contact : Feng Liang Email : fliang@life.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CSODI009DF05QP1.
FEATURES	Location/Qualifiers
source	1..1201
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone="CSODI009Y110"
	/tissue_type="PLACENTA COT 25-NORMALIZED"
	/note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN	
Query Match	50.0%; Score 1027.4; DB 13; Length 1201;
Best Local Similarity	95.6%; Pred. No. 4e-156;
Matches 1120; Conservative	9; Mismatches 35; Indels 7; Gaps 7;
QY	1 CCGGCGCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
DB	34 CCGGAATTTCCGGATCGCGCCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 93
QY	61 CGCGCGCTTCCTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
DB	94 CGCGCGCTTCCTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 152
QY	121 TCCCGGCGCGCGCTTCGCGCGCTTCGCGCGCTTCGCGCGCTTCGCGCGCTTCGCGCGCTTCGCGCG 180
DB	153 TCCCGGCGCGCGCTTCGCGCGCTTCGCGCGCTTCGCGCGCTTCGCGCGCTTCGCGCGCTTCGCGCG 212


```
421 ACTATCCCAACCTGTGTGCGAGCTGCGCGCGCGAGCGCGCTCGAGAGGCTGCACCGG 480
Db      ACCTATGCCCAACCTGTGTGCGAGCTGCGCGCGCGAGCGCGCTCGAGAGGCTGCACCGG 521
QY      CCGCGGCTCATCTGCTTGTGCGAGCGGAGCTGCGCGCGCGAGCGAGGAGATCCCAACAGT 540
Db      CCGCGGCTCATCTGCTTGTGCGAGCGGAGCTGCGCGCGCGAGCGAGGAGATCCCAACAGT 581
QY      TTGGGCCATAAATAAATACTTTATCGCGAGCGTGTGTGGAAGATGCGCCCTGCGGTGTT 600
Db      TTGGGCCATAAATAAATACTTTATCGCGAGCGTGTGTGGAAGATGCGCCCTGCGGTGTT 641
QY      CATATCGAATTTTTCGAGCTTCCGTTTCTAAACGAGAGGTGCGGTGCTAGTGG 660
Db      CATATCGAATTTTTCGAGCTTCCGTTTCTAAACGAGAGGTGCGGTGCTAGTGG 701
QY      TCTGGGTTATGTGTGCGAGATGAGCTGATCGTGAATAATGCGCACTGTGTGACCAAC 720
Db      TCTGGGTTATGTGTGCGAGATGAGCTGATCGTGAATAATGCGCACTGTGTGACCAAC 761
QY      AAGCACCGGTCCTGAGTGAAGACGGTGCACCTTACGAGCCAAATCAAGGAT 780
Db      AAGCACCGGTCCTGAGTGAAGACGGTGCACCTTACGAGCCAAATCAAGGAT 821
QY      GTGATGAGAAAGCAGACATCGCATCTCAATCAAAATTGACCAACAGGCGAAGTGCCTGTC 840
Db      GTGATGAGAAAGCAGACATCGCATCTCAATCAAAATTGACCAACAGGCGAAGTGCCTGTC 881
QY      CTGCTGCTTGGCGCTCTCAGAGCTGCGCGCGGAGAGTTCGTGTCGCGCATCGGAAGC 900
Db      CTGCTGCTTGGCGCTCTCAGAGCTGCGCGCGGAGAGTTCGTGTCGCGCATCGGAAGC 941
QY      CCGTTTCCCTTCAAAACACAGTACCAACCGGAGTGTGAGCAACACCGAGCGGCGG 960
Db      CCGTTTCCCTTCAAAACACAGTACCAACCGGAGTGTGAGCAACACCGAGCGGCGG 1001
QY      AAAGAGCTGGGCTCGGAACTCAGACATGAGTACATCCAGACCGAGCGCATCATCAAC 1020
Db      AAAGAGCTGGGCTCGGAACTCAGACATGAGTACATCCAGACCGAGCGCATCATCAAC 1060
QY      TATGGAATCGGGAGCGCGTTAGTAAACCTGGAACGTTGAGTGAATTAACACT 1080
Db      TATGGAATCGGGAGCGCGTTAGTAAACATGRCGSGTGAAGTGAATGGA-TAACACT 1118
QY      TTGAAGTGACA 1092
Db      TTGAAGTGAAA 1130

RESULT 7
BX423029/c
LOCUS      BX423029      1132 bp      mRNA      linear      EST 13-MAY-2003
DEFINITION BX423029 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone
            CS0DN003YF18 3-PRIME, mRNA sequence.
ACCESSION  BX423029
VERSION     BX423029.1  GI:30659434
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            Li,W.B., Gruber,C., Jessee,J. and Polyes,D.
            Full-length cDNA libraries and normalization
            Unpublished (2001)
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 5631.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
```

```
cgi-bin/cluster.cgi?seq=CS0DN003DC09NP1&cluster=5631.f. Contact :
Peng Liang Email : filiang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DN003DC09NP1.
Location/Qualifiers
1. 1132
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DN003YF18"
/tissue_type="ADULT BRAIN"
/dev_stage="adult"
/clone_lib="Homo sapiens ADULT BRAIN"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

FEATURES
source
Query Match 49.5%; Score 1016.8; DB 13; Length 1132;
Best Local Similarity 97.6%; Pred. No. 2e-154;
Matches 1039; Conservative 7; Mismatches 16; Indels 3; Gaps 2;

QY      945  CACCCAGCGAGCGCGCAAGAGCTGGGCTCCGCACTCAGACATCGACTACATCCAGAC 1004
Db      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY      1062  CCMCMGCSAAGCGGCAAGAGCTGGG--CTCCGCACTCAGACAT-GAATACATCCAGAC 1006
Db      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY      1005  CGAGCCCATCATCAACTATGGAACCTCGGAGGCGCGGTTAGTAAACCTGGACGGTGAAGT 1064
Db      1005  CGAGCCCATCATCAACTATGGAACCTCGGAGGCGCGGTTAGTAAACCTGGACGGTGAAGT 946
QY      1065  GATTGGAAATTAACACTTTGAAAGTGACAGCTGGAATCTCTTTGCAATCCATCTGATAA 1124
Db      945  GATTGGAAATTAACACTTTGAAAGTGACAGCTGGAATCTCTTTGCAATCCATCTGATAA 886
QY      1125  GATTAAAAAGTTCTCTCAGCGAGTCCCATGACCGACGAGGCGCAAGGAAAGCATCAACAA 1184
Db      886  GATTAAAAAGTTCTCTCAGCGAGTCCCATGACCGACGAGGCGCAAGGAAAGCATCAACAA 826
QY      1185  GAAGAGTATATTGTTATCGGATGATGTCACCTCAGCTCCAGCAAGCCAAAGAGCTGAA 1244
Db      825  GAAGAGTATATTGTTATCGGATGATGTCACCTCAGCTCCAGCAAGCCAAAGAGCTGAA 766
QY      1245  GGACCGGACCGGGACTTCCCGCAATGATGTCAGGAGCGTATATAATTGAAGTAATTC 1304
Db      765  GGACCGGACCGGGACTTCCCGCAATGATGTCAGGAGCGTATATAATTGAAGTAATTC 706
QY      1305  TGATACCCCGAGAGAGCTGGTCTCAAGGAACGAGCTCATATATCATGATCAATGG 1364
Db      705  TGATACCCCGAGAGAGCTGGTGGTCTCAAGGAACGAGCTCATATATCATGATCAATGG 646
QY      1365  ACAGTCCGTGTCTCCGCAATGATGTCAGCGAGCTCATTAAGGAAAGGAAAGCAACCTGAA 1424
Db      645  ACAGTCCGTGTCTCCGCAATGATGTCAGCGAGCTCATTAAGGAAAGGAAAGCAACCTGAA 586
QY      1425  CATGGTGTCCGAGGCGGTAATGAGATATCATGATCAGATGATCCCGCAAGAAATGA 1484
Db      585  CATGGTGTCCGAGGCGGTAATGAGATATCATGATCAGATGATCCCGCAAGAAATGA 526
QY      1485  CCCATAGGCGAGGAGTGAAGTGGACTTCATGTTTCCCTCAAGACTCTCCCGTGGATGA 1544
Db      525  CCCATAGGCGAGGAGTGAAGTGGACTTCATGTTTCCCTCAAGACTCTCCCGTGGATGA 466
QY      1545  CGGATGAGACTCTGGGCTGGATAGGACACTCAAGACTTTTGACTGCCATTTGTT 1604
Db      465  CGGATGAGACTCTGGGCTGGATAGGACACTCAAGACTTTTGACTGCCATTTGTT 406
QY      1605  TGTTCAGTGGAGACTCCCTGGCCAAACAGAAATCTTTCTTGATGTTTCAGGCAACAAA 1664
Db      405  TGTTCAGTGGAGACTCCCTGGCCAAACAGAAATCTTTCTTGATGTTTCAGGCAACAAA 346
QY      1665  TGTATGTTGAGATCCCGAGGAGGCTCTGCCCTTCTGTCATCTATGATGATGAGTGT 1724
```


Db	345	TGTAATGTTGCAGATCCGACAGGAGAGCTCTCCCTTCTGTATCTATGATATGCAGTGT	286
Qy	1725	GCCTTTCTTTCAGCTTGGGCCATTCTTGCTTTAGACAGTCAGCAATTTGCTCCTCCTTT	1784
Db	285	GCCTTTCTTTCAGCTTGGGCCATTCTTGCTTTAGACAGTCAGCAATTTGCTCCTCCTTT	226
Qy	1785	AACGTAGTCATCATCTTAGTCCAACTAAATGCAGTCGATACAAATGCGTAGATGAAGAAGC	1844
Db	225	AACGTAGTCATCATCTTAGTCCAACTAAATGCAGTCGATACAAATGCGTAGATGAAGAAGC	166
Qy	1845	CCACGGGACCCAGGATGGAGCTGGTCGTGTTTGTGCTTTCTCCAAAGTCAGCACCCAAA	1904
Db	165	CCACGGGACCCAGGATGGAGCTGGTCGTGTTTGTGCTTTCTCCAAAGTCAGCACCCAAA	106
Qy	1905	GGTCAATGCACAGAGACCCCGGTGGGTGAGCGCTGGCTTCTCAACGCGCGAAGTTGCC	1964
Db	105	GGTCAATGCACAGAGACCCCGGTGGGTGAGCGCTGGCTTCTCAACGCGCGAAGTTGCC	46
Qy	1965	TCCTTTAGGAATCTCTTTGGAAATGGGACGACGACTCTGAGT	2009
Db	45	TCCTTTAGGAATCCTNNGGAANNGGACGACGANGAGACTOMK 1	
RESULT 8			
BX381058/c			
LOCUS			
DEFINITION			
BX381058 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA			
clone CS0D1060YC24 3-PRIME, mRNA sequence.			
ACCESSION			
BX381058			
KEYWORDS			
EST.			
SOURCE			
Homo sapiens (human)			
ORGANISM			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
Contact: Genoscope			
Genoscope - Centre National de Sequencage			
BP 131 91006 EVRY cedex - France			
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr			
Library was constructed by Life Technologies, a division of			
Invitrogen. This sequence belongs to sequence cluster 5631.f For			
more information about this cluster, see			
http://www.genoscope.cns.fr/			
cgi-bin/cluster.cgi?seq=CS0D1060B12NP1&cluster=5631.f. Contact :			
Feng Liang Email: fliang@lifetech.com URL :			
http://fulllength.invitrogen.com/ invitrogen Corporation 1600			
Paradise Avenue Genoscope sequence ID : CS0D1060B12NP1.			
Location/Qualifiers			
1. .1201			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="CS0D1060YC24"			
/tissue_type="PLACENTA COT 25-NORMALIZED"			
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"			
/notes="1st strand cDNA was primed with a NotI-oligo(dT)			
primer. Five prime end enriched, double-strand cDNA was			
digested with Not I and cloned into the Not I and EcoR V			
sites of the pCMVSPORT 6 vector. Library was normalized."			
FEATURES			
source			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 1054; Conservative			
49.3%; Score 1013.6; DB 13; Length 1201;			
96.3%; Pred. No. 6.7e-154;			
10; Mismatches 27; Indels			
3; Gaps			
Qy	911	TTCAAAACACAGTCACACCGGGATCTGTAGCACCACCCGAGCGGCGAAAGAGCTGG	970
Db	1106	TTCTTTAAACACATCCWCWCGGGATCTGTGACACCCACCCGAGCGGCGAAAGAGTGG	1047

```

RESULT 9
BX423083/c
LOCUS
DEFINITION BX423083 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone
CS0DN004YC10 3-PRIME, mRNA sequence.
ACCESSION BX423083
VERSION BX423083.1 GI:30643301
KEYWORDS EST
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1090)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5631.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DN004BB05NP1&cluster=5631.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue Genoscope sequence ID : CS0DN004BB05NP1.
FEATURES
Location/Qualifiers
1..1090
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DN004YC10"
/tissue_type="ADULT BRAIN"
/dev_stages="adult"
/clone_lib="Homo sapiens ADULT BRAIN"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."
ORIGIN
Query Match 48.4%; Score 994; DB 13; Length 1090;
Best Local Similarity 98.8%; Pred. No. 9.9e-151;
Matches 1019; Conservative 2; Mismatches 8; Indels 2; Gaps 2;
QY 944 CCACCCAGCGAGCGGCGAAAGAGCTGGGGCTCCGCAACTCAGACATGGACTACATCCAGA 1003
DB :
1074 MCACCCAGCGAGGC-GCAAAGAGCTGGGCTCCGCAAYCTCAGACAT-GACTACATCCAGA 1017
QY 1004 CCGACGCCATCATCAACTATGGAACCTGGGAGGCCGTTAGTAAACCTGGACGGTGAAG 1063
DB 1016 CCGACGCCATCATCACTATGGAACCTGGGAGGCCGTTAGTAAACCTGGACGGTGAAG 957
QY 1064 TGAATTGGAATTAACACTTTGGAAGTGACAGCTGGAATCTCTTTGCAATCCCATCTGATA 1123
DB 956 TGAATTGGAATTAACACTTTGGAAGTGACAGCTGGAATCTCTTTGCAATCCCATCTGATA 897
QY 1124 AGATTAAGTCTCTCAGGAGTCCCATGACGACGACGCCCAAGGAAAGCCATCACCA 1183
DB 896 AGATTAAGTCTCTCAGGAGTCCCATGACGACGACGCCCAAGGAAAGCCATCACCA 837
QY 1184 AGAAGAGTATATTGGTATCCGAATGATGTCATCACTACGTCGCCAGCAAGCAAGAGCTGA 1243
DB 836 AGAAGAGTATATTGGTATCCGAATGATGTCATCACTACGTCGCCAGCAAGCAAGAGCTGA 777
QY 1244 AGGACCGGACCGGAGCTCCGAGAGTATCTCAGGAGGTATATAATTGAAGTAATTC 1303
DB 776 AGGACCGGACCGGAGCTCCGAGAGTATCTCAGGAGGTATATAATTGAAGTAATTC 717

```

```

QY 1304 CTGATACCCAGCAGAAAGCTGGTGTCTCAAGGAAAAACGACGTCATAATCAGCATCAATG 1363
DB 716 CTGATACCCAGCAGAAAGCTGGTGTCTCAAGGAAAAACGACGTCATAATCAGCATCAATG 657
QY 1364 GACAGTCCGTGGTCTCCGCCAATGATGTCAGGAGCGTCAATTAAGGGAAGAACCCCTGA 1423
DB 656 GACAGTCCGTGGTCTCCGCCAATGATGTCAGGAGCGTCAATTAAGGGAAGAACCCCTGA 597
QY 1424 ACATGGTGGTCCGAGGGGTAAATGAAGATATCATGATCACAGTGTATCCGGAAGAAATTG 1483
DB 596 ACATGGTGGTCCGAGGGGTAAATGAAGATATCATGATCACAGTGTATCCGGAAGAAATTG 537
QY 1484 ACCCATAGGAGAGGATGAGCTGAGCTTATGTTTCCCTCAAAGACTCTCCCGTGGATG 1543
DB 536 ACCCATAGGAGAGGATGAGCTGAGCTTATGTTTCCCTCAAAGACTCTCCCGTGGATG 477
QY 1544 ACGGATCAGGACTCTGGGCTCTGGAATAGGACACTCAAGACTTTTGACTGCCATTTTGT 1603
DB 476 ACGGATCAGGACTCTGGGCTCTGGAATAGGACACTCAAGACTTTTGACTGCCATTTTGT 417
QY 1604 TTGTTCACTGGAGACTCCCTGGCCAAACAGAAATCCCTTTTGATAGTTTGAGGCAAAACAA 1663
DB 416 TTGTTCACTGGAGACTCCCTGGCCAAACAGAAATCCCTTTTGATAGTTTGAGGCAAAACAA 357
QY 1664 ATGTAATGTTGCAGATCGCAGGAGGAGCTCTGCCCTTCTGTATCTATGATGAGTG 1723
DB 356 ATGTAATGTTGCAGATCGCAGGAGGAGCTCTGCCCTTCTGTATCTATGATGAGTG 297
QY 1724 TGTCTTTTCTGCGAGCTTGGGCCAATCTTGCTTAGACAGTCAAGCTTTGCTCTCTCTT 1783
DB 296 TGTCTTTTCTGCGAGCTTGGGCCAATCTTGCTTAGACAGTCAAGCTTTGCTCTCTCTT 237
QY 1784 TAACCTGAGTCATCATCTTAGTCCCAACTAATGACATGATACAAATGCTAGATAGAAAG 1843
DB 236 TAACCTGAGTCATCATCTTAGTCCCAACTAATGACATGATACAAATGCTAGATAGAAAG 177
QY 1844 CCCACCGGAGCCAGGATGGGACTGCTGTGTTTGTGCTTTCTCAAGTCAGACCCCAA 1903
DB 176 CCCACCGGAGCCAGGATGGGACTGCTGTGTTTGTGCTTTCTCAAGTCAGACCCCAA 117
QY 1904 AGGTCAATGACAGAGACCCCGGGTGGGTGAGCGTGGCTTCTCAACGCGCGAAGTGC 1963
DB 116 AGGTCAATGACAGAGACCCCGGGTGGGTGAGCGTGGCTTCTCAACGCGCGAAGTGC 57
QY 1964 CTCTTTTAGGA 1974
DB 56 STCTTTTAGGA 46
RESULT 10
BX339704/c
LOCUS
DEFINITION BX339704 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0DI075YN06 3-PRIME, mRNA sequence.
ACCESSION BX339704
VERSION BX339704.1 GI:30341839
KEYWORDS EST
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5631.f For
more information about this cluster, see
http://www.genoscope.cns.fr/

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 10, 2004, 08:38:17 ; Search time 2825.63 Seconds
(without alignments)
17011.249 Million cell updates/sec

Title: US-10-084-817-37

Perfect score: 1109

Sequence: 1 cctgaccaggagacactgg.....acanaattccaccctatccg 1109

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

1: gb_ba:

2: gb_htg:

3: gb_in:

4: gb_ov:

5: gb_pat:

6: gb_ph:

7: gb_pl:

8: gb_ro:

9: gb_sy:

10: gb_vt:

11: gb_vt:

12: gb_vt:

13: gb_vt:

14: gb_vt:

15: gb_vt:

16: gb_vt:

17: gb_vt:

18: gb_vt:

19: gb_vt:

20: gb_vt:

21: gb_vt:

22: gb_vt:

23: gb_vt:

24: gb_vt:

25: gb_vt:

26: gb_vt:

27: gb_vt:

28: gb_vt:

29: gb_vt:

30: gb_vt:

31: gb_vt:

32: gb_vt:

33: gb_vt:

34: gb_vt:

35: gb_vt:

36: gb_vt:

37: gb_vt:

38: gb_vt:

39: gb_vt:

40: gb_vt:

41: gb_vt:

42: gb_vt:

43: gb_vt:

44: gb_vt:

45: gb_vt:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	802	72.3	851	6	AX014337	AX014337 Sequence
2	802	72.3	851	6	BD222216	BD222216 Human nuc
3	699.8	63.1	1095	6	AX281854	AX281854 Sequence
4	696.2	62.8	853	6	AX397528	AX397528 Sequence
5	696.2	62.8	853	6	AX302539	AX302539 Sequence
6	696.2	62.8	853	6	AX409461	AX409461 Sequence
7	696.2	62.8	853	9	HUM927A	J04164 Human inter
8	664.6	59.9	683	9	BC000897	BC000897 Homo sapi
9	624.4	56.3	647	9	HSR9A927	X84958 H.sapiens m
10	378	34.1	378	9	BT007173	BT007173 Homo sapi
11	376.4	33.9	378	12	BT007946	BT007946 Synthetic
12	362	32.6	156819	2	AP006283	AP006283 Homo sapi
13	362	32.6	175416	2	AP006286	AP006286 Homo sapi
14	362	32.6	187160	9	AC138230	AC138230 Homo sapi
15	360.4	32.5	143779	9	AC136475	AC136475 Homo sapi
16	351	31.7	352	6	AR273461	AR273461 Sequence
17	351	31.7	352	6	AR277042	AR277042 Sequence
18	351	31.7	352	6	AR407317	AR407317 Sequence
19	351	31.7	352	6	AX368494	AX368494 Sequence
20	340.8	30.7	622	4	AF272042	AF272042 Bos tauru
21	335.8	30.3	764	6	AR148104	AR148104 Sequence
22	335.8	30.3	764	6	AR350926	AR350926 Sequence
23	335.8	30.3	764	6	AR433302	AR433302 Sequence
24	335.8	30.3	764	6	AX316946	AX316946 Sequence
25	335.8	30.3	764	6	BD084479	BD084479 Compositi
26	302.6	27.3	695	6	AX281855	AX281855 Sequence
27	296.4	26.7	355	9	HSIFNIN2	X02491 Human inter
28	289.6	26.1	637	9	BC006794	BC006794 Homo sapi
29	286	25.8	316	6	AX397529	AX397529 Sequence
30	285.4	25.7	122351	9	AC023157	AC023157 Homo sapi
31	282.4	25.5	621	6	BD190298	BD190298 Elk1 phos
32	281.8	25.4	808	6	AR380734	AR380734 Sequence
33	281.8	25.4	808	6	BD190297	BD190297 Elk1 phos
34	281.8	25.4	808	6	HS18U	X57352 Human 1-8U
35	280.2	25.3	222085	2	AC144988	AC144988 Gorilla g
36	278.6	25.1	749	9	BC008417	BC008417 Homo sapi
37	277.8	25.0	645	9	BC022439	BC022439 Homo sapi
38	277.2	25.0	905	6	BD190334	BD190334 Elk1 phos
39	275.6	24.9	905	6	BD190333	BD190333 Elk1 phos
40	275.4	24.8	642	6	BD237072	BD237072 Compounds
41	275.4	24.8	642	6	AR225472	AR225472 Sequence
42	275.4	24.8	642	6	AX321588	AX321588 Sequence
43	274	24.7	905	6	AX337050	AX337050 Sequence
44	274	24.7	905	6	BD190295	BD190295 Elk1 phos
45	274	24.7	905	9	HS18D	X57351 Human 1-8D

ALIGNMENTS

RESULT 1	AX014337	AX014337	851 bp	DNA	linear	PAT 07-SEP-2000
AX014337	LOCUS	Sequence 46 from Patent WO9954353.				
DEFINITION	AX014337	AX014337				
ACCESSION	AX014337	AX014337				
VERSION	AX014337.1	GI:10040691				
KEYWORDS		Homo sapiens (human)				
SOURCE		Homo sapiens				
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
REFERENCE		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS		Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and				
TITLE		Pilarsky,C.				
		Human nucleic acid sequences of normal uterus tissue				

JOURNAL Patent: WO 9954353-A 46 28-OCT-1999;
 SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
 BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN
 (DE); PILARSKY CHRISTIAN (DE)

FEATURES
 source
 Location/Qualifiers
 1. 851
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Query Match 72.3%; Score 802; DB 6; Length 851;
 Best Local Similarity 98.6%; Pred. No. 1.3e-180;
 Matches 830; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

QY 1 CTTGACAGAGACACTGGAGGTTTTCAGGAGAGAGGCTCGGACAAAGCCGACAGAGAGGACTGCAGC 60
 DB 2 CTTGACAGAGACACTGGAGGTTTTCAGGAGAGAGGCTCGGACAAAGCCGACAGAGAGGACTGCAGC 61
 QY 61 CTGAGGAAAGAGCAAGGATTTTCAGGAGAGAGGCTCGGACAAAGCCGACAGAGAGGACTGCAGC 120
 DB 62 CTGAGGAAAGAGCAAGGATTTTCAGGAGAGAGGCTCGGACAAAGCCGACAGAGAGGACTGCAGC 121
 QY 121 CTTAAGAGAAATACACACTTCCTGAGAACTGAAACCGACAGAGGAAAGGAGTCTCACTGA 180
 DB 122 CTTAAGAGAAATACACACTTCCTGAGAACTGAAACCGACAGAGGAAAGGAGTCTCACTGA 180
 QY 181 GCACCGTCCAGCATCCGACACACAGCGGCTTCGCTCCAGCGGAAACACACACTT 240
 DB 181 GCACCGTCCAGCATCCGACACACAGCGGCTTCGCTCCAGCGGAAACACACACTT 240
 QY 241 CTCACACCTTCACTCAACACTTCCTTCCCAAGCCAGAGAGATGTCACAGAGGAAACATG 300
 DB 241 CTCACACCTTCACTCAACACTTCCTTCCCAAGCCAGAGAGATGTCACAGAGGAAACATG 300
 QY 301 AGGTGGCTGTGTGGGGGCAACCCCGACGACCACTTCCTTCCCAAGCCAGAGAGATGTCACAG 360
 DB 301 AGGTGGCTGTGTGGGGGCAACCCCGACGACCACTTCCTTCCCAAGCCAGAGAGATGTCACAG 360

QY 421 TGAAGTGTGTGTGTGGGCTTCATAGCATTCGCTTCTCCAGGCTCCAGGACAGGA 480
 DB 421 TGAAGTGTGTGTGTGGGCTTCATAGCATTCGCTTCTCCAGGCTCCAGGACAGGA 480
 QY 481 AGATGTTGGGAGCTGACCGGGGCGGAGGCTATGCTCCAGCGGCAAGTGCCTGACGA 540
 DB 481 AGATGTTGGGAGCTGACCGGGGCGGAGGCTATGCTCCAGCGGCAAGTGCCTGACGA 540
 QY 541 TCTGGGCTGTATCTGGGCTTCCTATGACCATTCGATTCATCTCTGTTACTGGTATTCG 600
 DB 541 TCTGGGCTGTATCTGGGCTTCCTATGACCATTCGATTCATCTCTGTTACTGGTATTCG 600
 QY 601 GCTGTGACAGTCACTATATTTAGTACAGTAATACAGGAAACCGGGTTACTAGT 660
 DB 601 GCTGTGACAGTCACTATATTTAGTACAGTAATACAGGAAACCGGGTTACTAGT 660
 QY 661 AGCGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGTGGCCCTGCGAGCTGGG 720
 DB 661 AGCGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGTGGCCCTGCGAGCTGGG 720
 QY 721 GCTGTGTGCCCTTGCCCTTGCTGCTCCCTAGATACAGCAGTTTATACCCACACCT 780
 DB 721 GCTGTGTGCCCTTGCCCTTGCTGCTCCCTAGATACAGCAGTTTATACCCACACCT 780
 QY 781 GTCACACTGACATTCATATAAGTG-ACGTGCTTGTGAAAAAATAAATAAATACCCGA 839
 DB 781 GTCACACTGACATTCATATAAGTGACGTGCTTGTGAAAAAATAAATAAATACCCGA 840
 QY 840 GG 841

Db 841 GG 842

RESULT 2
 BD222216
 LOCUS
 DEFINITION Human nucleic acid sequence originating in normal uterine tissue.
 ACCESSION BD222216
 VERSION 1 GI:33031986
 KEYWORDS UP 2002512017-A/45.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 1 (bases 1 to 851)
 Specht, T., Hinzmann, B., Schmitt, A., Pilarczyk, C., Dahl, E. and
 Rosenthal, A.
 Human nucleic acid sequence originating in normal uterine tissue
 Patent: JP 2002512017-A 45 23-APR-2002;
 METAGEN GESELLSCHAFT FUER GENOM FORSCHUNG MBH
 OS Homo sapiens (human)
 PN JP 2002512017-A/45
 PD 23-APR-2002
 PF 15-APR-1999 JP 2000544691
 PR 17-APR-1998 DE 198 17 945 4
 PT THOMAS SPECHT, BERND HINZMANN, ARMIN SCHMITT, CHRISTIAN PILARSKY,
 EDGAR DAHL,
 PI ANDRE ROSENTHAL
 PC C12N15/09, A61K38/00, A61K48/00, C07K14/47, C07K16/18,
 C12N1/19,
 PC C12N1/21, C12N5/00, C12P21/08, C12Q1/68, G01N33/68, C12N15/00, PC
 A61K37/02,
 PC C12N5/00
 CC Human nucleic acid sequence originating in normal uterine CC
 tissue

FEATURES
 source
 1. 851
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

ORIGIN
 Query Match 72.3%; Score 802; DB 6; Length 851;
 Best Local Similarity 98.6%; Pred. No. 1.3e-180;
 Matches 830; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

QY 1 CTTGACAGAGACACTGGAGGTTTTCAGGAGAGAGGCTCGGACAAAGCCGACAGAGAGGACTGCAGC 60
 DB 2 CTTGACAGAGACACTGGAGGTTTTCAGGAGAGAGGCTCGGACAAAGCCGACAGAGAGGACTGCAGC 61
 QY 61 CTGAGGAAAGAGCAAGGATTTTCAGGAGAGAGGCTCGGACAAAGCCGACAGAGAGGACTGCAGC 120
 DB 62 CTGAGGAAAGAGCAAGGATTTTCAGGAGAGAGGCTCGGACAAAGCCGACAGAGAGGACTGCAGC 121
 QY 121 CTTAAGAGAAATACACACTTCCTGAGAACTGAAACCGACAGAGGAAAGGAGTCTCACTGA 180
 DB 122 CTTAAGAGAAATACACACTTCCTGAGAACTGAAACCGACAGAGGAAAGGAGTCTCACTGA 180
 QY 181 GCACCGTCCAGCATCCGACACACAGCGGCTTCGCTCCAGCGGAAACACACACTT 240
 DB 181 GCACCGTCCAGCATCCGACACACAGCGGCTTCGCTCCAGCGGAAACACACACTT 240
 QY 241 CTCACACCTTCACTCAACACTTCCTTCCCAAGCCAGAGAGATGTCACAGAGGAAACATG 300
 DB 241 CTCACACCTTCACTCAACACTTCCTTCCCAAGCCAGAGAGATGTCACAGAGGAAACATG 300
 QY 301 AGGTGGCTGTGTGGGGGCAACCCCGACGACCACTTCCTTCCCAAGCCAGAGAGATGTCACAG 360
 DB 301 AGGTGGCTGTGTGGGGGCAACCCCGACGACCACTTCCTTCCCAAGCCAGAGAGATGTCACAG 360

QY 361 TCCACAGCAGACCTCGTGGCCGACCATGCTGCTGGTCCCTCTTCAACACCTCTTCT 420
 Db 361 TCCACAGCAGACCTCGTGGCCGACCATGCTGCTGGTCCCTCTTCAACACCTCTTCT 420
 QY 421 TGAACCTGGTGTCTGCTGGGCTTTCATAGCATTCGCTACTCTCGTGAAGTCTAGGACAGGA 480
 Db 421 TGAACCTGGTGTCTGCTGGGCTTTCATAGCATTCGCTACTCTCGTGAAGTCTAGGACAGGA 480
 QY 481 AGATGGTGGCAGCTGACCGGGGCCAGCCTATGCTCCACCGCAAGTCCCTGAACA 540
 Db 481 AGATGGTGGCAGCTGACCGGGGCCAGCCTATGCTCCACCGCAAGTCCCTGAACA 540
 QY 541 TCTGGGCCCTGATCTTCTGGGCATCTCATGACCATGGATTTCATCTGTTACTGTTATTCG 600
 Db 541 TCTGGGCCCTGATCTTCTGGGCATCTCATGACCATGGATTTCATCTGTTACTGTTATTCG 600
 QY 601 GCTCTGTGACAGTCTACCATATTTATGTTACAGATAATACAGGAAACCGGGTTACTAGT 660
 Db 601 GCTCTGTGACAGTCTACCATATTTATGTTACAGATAATACAGGAAACCGGGTTACTAGT 660
 QY 661 AGCGGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTGGCCCTGCACGCTGGG 720
 Db 661 AGCGGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTGGCCCTGCACGCTGGG 720
 QY 721 GCTGTGCCCCCTGCCCCCTGGTCTCCCTAGATACAGAGTATTATACCCACACCT 780
 Db 721 GCTGTGCCCCCTGCCCCCTGGTCTCCCTAGATACAGAGTATTATACCCACACCT 780
 QY 781 GTCACACTGACATTCATTAAGAGT-ACGTGCTTGTGAAAAAATAAATACCCGA 839
 Db 781 GTCACACTGACATTCATTAAGAGT-ACGTGCTTGTGAAAAAATAAATACCCGA 839
 QY 840 GG 841
 Db 841 GG 842

RESULT 3
 AX281854
 LOCUS 1095 bp DNA linear PAT 02-NOV-2001
 DEFINITION Sequence 263 from Patent WO0177389.
 ACCESSION AX281854
 VERSION AX281854.1 GI:16609105
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Shiffman, D., Somogyi, R., Lawn, R., Seilhamer, J.J., Porter, G.J.,
 Mikita, T. and Tai, J.
 TITLE Genes expressed in foam cell differentiation
 JOURNAL Patent: WO 0177389-A 263 18-OCT-2001;
 Incyte Genomics, Inc. (US)
 FEATURES Location/Qualifiers
 source
 1. 1095
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="Incyte ID No: 988231.7"

ORIGIN
 Query Match 63.1%; Score 699.8; DB 6; Length 1095;
 Best Local Similarity 99.6%; Pred. No. 3.1e-156;
 Matches 712; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 106 AGCAGGAATAGAACTTAAGAGAAATACACATTTCTGAGAAATCTGAGAACTGAAACGACAGGGGAA 165
 Db 381 AGCAGGAATAGAACTTAAGAGAAATACACATTTCTGAGAAATCTGAGAACTGAAACGACAGGGGAA 440
 QY 166 AGGAGTCTACCTGAGCAGCTCCAGCATCCGGACACACAGCGGCCCTTGGCTCCAG 225
 Db 441 AGGAGTCTACCTGAGCAGCTCCAGCATCCGGACACACAGCGGCCCTTGGCTCCAG 500

QY 226 CAGAAAAACACATCTCTCAAAACCTTCACTCAACACTTCTCTCCCAACCCAGAGATGC 285
 Db 501 CAGAAAAACACATCTCTCAAAACCTTCACTCAACACTTCTCTCCCAACCCAGAGATGC 550
 QY 286 ACAAGGAGGAACATGAGTGGCTGTGTGGGGGACCCGCCAGCAGCACCATTCTTCCAAGGT 345
 Db 561 ACAAGGAGGAACATGAGTGGCTGTGTGGGGGACCCGCCAGCAGCACCATTCTTCCAAGGT 620
 QY 346 CACCCCTGATCAATCCACAGCAGACCTCTGTCGCCGACCATGCTGCTGGTCCCTGT 405
 Db 621 CACCCCTGATCAATCCACAGCAGACCTCTGTCGCCGACCATGCTGCTGGTCCCTGT 680
 QY 406 TCAACACCCCTCTTCTTGAACCTGGTGTCTGTGGGCTTTCATAGCATTCGCTACTCTCGTGA 465
 Db 681 TCAACACCCCTCTTCTTGAACCTGGTGTCTGTGGGCTTTCATAGCATTCGCTACTCTCGTGA 740
 QY 466 AGTCTAGGACAGAGATGTTGGCGACGTGACCGGGGCCAGCCCTATGCTTCCACCG 525
 Db 741 AGTCTAGGACAGAGATGTTGGCGACGTGACCGGGGCCAGCCCTATGCTTCCACCG 800
 QY 526 CCAAGTGCCTGAACATCTGGGCCCTGATTCTGGGCATCTCATGACCATTGGATTTCATCC 585
 Db 801 CCAAGTGCCTGAACATCTGGGCCCTGATTCTGGGCATCTCATGACCATTGGATTTCATCC 860
 QY 586 TGTACTGGTATTGGCTCTGTGACAGTCTACCATATTATTTACAGATAATACAGGAA 645
 Db 861 TGTACTGGTATTGGCTCTGTGACAGTCTACCATATTATTTACAGATAATACAGGAA 920
 QY 646 AACGGGTTACTAGTACCGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTG 705
 Db 921 AACGGGTTACTAGTACCGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTG 980
 QY 706 GCCTGACAGCTGGGGCTGTTGGCCCTGCTGGCCCTGCTGGCCCTAGATACAGAGTT 765
 Db 981 GCCTGACAGCTGGGGCTGTTGGCCCTGCTGGCCCTGCTGGCCCTAGATACAGAGTT 1040
 QY 766 TATACCCACACACTGTCTACACTGACATTCATTAAGAGT-ACGTGCTTGTGAAA 819
 Db 1041 TATACCCACACACTGTCTACAGTGTCAATCAATAAGTGACAGCTTGTGTGAAA 1095

RESULT 4
 AX397528
 LOCUS 842 bp DNA linear PAT 18-MAY-2002
 DEFINITION Sequence 1743 from Patent WO0212328.
 ACCESSION AX397528
 VERSION AX397528.1 GI:21068275
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS King, G.E., Meagher, M.J., Xu, J. and Secrist, H.
 TITLE Compositions and methods for the therapy and diagnosis of colon cancer
 JOURNAL Patent: WO 0212328-A 1743 14-FEB-2002;
 CORIXA CORPORATION (US)
 FEATURES Location/Qualifiers
 source
 1. 842
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Query Match 62.8%; Score 696.2; DB 6; Length 842;
 Best Local Similarity 99.4%; Pred. No. 2.3e-155;
 Matches 709; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 106 AGCAGGAATAGAACTTAAGAGAAATACACATTTCTGAGAAATCTGAGAACTGAAACGACAGGGGAA 165
 Db 130 AGCAGGAATAGAACTTAAGAGAAATACACATTTCTGAGAAATCTGAGAACTGAAACGACAGGGGAA 189

166 AGGAGGTCTCAGTGGACCGTCCAGCATCCGGACACACAGCGGCCCTTCGCTCCACG 225
 Db
 190 AGGAGGTCTCAGTGGACCGTCCAGCATCCGGACACACAGCGGCCCTTCGCTCCACG 249
 Qy
 226 CAGAAACACACATCTCAAAACCTTCACTCAACATCTCTTCCCAAGCCAGAGATGC 285
 Db
 250 CAGAAACACACATCTCAAAACCTTCACTCAACATCTCTTCCCAAGCCAGAGATGC 309
 Qy
 286 ACAAGGAGGAACATGAGTGGCTGTCTGGGGCCACCCCGCCAGCAGCATCTTCCAAAGT 345
 Db
 310 ACAAGGAGGAACATGAGTGGCTGTCTGGGGCCACCCCGCCAGCAGCATCTTCCAAAGT 369
 Qy
 346 CCACCGTGATCAACATCCACAGCGAGACTCCGTGCGCCGACCATGTCGTCGCTCCCTGT 405
 Db
 370 CCACCGTGATCAACATCCACAGCGAGACTCCGTGCGCCGACCATGTCGTCGCTCCCTGT 429
 Qy
 406 TCAACACCCCTCTTCTGAACTGGTGTCTCTGGGCTTCATAGCATTCGCGCTACTCCGTGA 465
 Db
 430 TCAACACCCCTCTTCTGAACTGGTGTCTCTGGGCTTCATAGCATTCGCGCTACTCCGTGA 489
 Qy
 466 AGTCTAGGACAGGAAGTGTGGCGAGTACCGGGGCCCGAGGCTATGCTCCACCG 525
 Db
 490 AGTCTAGGACAGGAAGTGTGGCGAGTACCGGGGCCCGAGGCTATGCTCCACCG 549
 Qy
 526 CCAAGTCCCTGAACATCTGGGCCCTGATCTGGGCATCTCATGACCATTTGGATTTCATCC 585
 Db
 550 CCAAGTCCCTGAACATCTGGGCCCTGATCTGGGCATCTCATGACCATTTGGATTTCATCC 609
 Qy
 586 TGTACTGGTATTCGGCTCTGTGACAGTCTACCATATTATTTACAGATAAATACAGGAAA 645
 Db
 610 TGTACTGGTATTCGGCTCTGTGACAGTCTACCATATTATTTACAGATAAATACAGGAAA 669
 Qy
 646 AACGGGTTACTAGTAGCGGCCATAGCTGCAACTTTGGCTCCACTCCAGTGCATGCTG 705
 Db
 670 AACGGGTTACTAGTAGCGGCCATAGCTGCAACTTTGGCTCCACTCCAGTGCATGCTG 729
 Qy
 706 GCCCTGCACGCTGGGGCTGTTCGCCCTTGGCTCCCTTGGCTCCCTTAGATACAGCAGTT 765
 Db
 730 GCCCTGCACGCTGGGGCTGTTCGCCCTTGGCTCCCTTGGCTCCCTTAGATACAGCAGTT 789
 Qy
 766 TATACCCACACACCTGTCTACAGTGTCTATCAATAAAGTG-ACGTGCTTGTGA 817
 Db
 790 TATACCCACACACCTGTCTACAGTGTCTATCAATAAAGTG-ACGTGCTTGTGA 842

RESULT 5
 AX302539
 LOCUS AX302539
 DEFINITION Sequence 57 from Patent WO0175177.
 ACCESSION AX302539
 VERSION AX302539.1 GI:17383079
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
 1
 Morin,P.J., Sherman-Baust,C.A., Pizer,E.S. and Hough,C.D.
 Tumor markers in ovarian cancer
 Patent: WO 0175177-A 57 11-OCT-2001;
 THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (US)
 LOCATION/Qualifiers
 1..853
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Query Match 62.8%; Score 696.2; DB 6; Length 853;
 Best Local Similarity 99.4%; Pred. No. 2.3e-155;
 Matches 709; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

106 AGCAGGAAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAA 165
 Db
 141 AGCAGGAAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAA 200
 Qy
 166 AGGAGGTCTCAGTGGACCGTCCAGCATCCGGACACACAGCGGCCCTTCGCTCCACG 225
 Db
 201 AGGAGGTCTCAGTGGACCGTCCAGCATCCGGACACACAGCGGCCCTTCGCTCCACG 260
 Qy
 226 CAGAAACACACATCTCAAAACCTTCACTCAACATCTCTTCCCAAGCCAGAGATGC 285
 Db
 261 CAGAAACACACATCTCAAAACCTTCACTCAACATCTCTTCCCAAGCCAGAGATGC 320
 Qy
 286 ACAAGGAGGAACATGAGTGGCTGTCTGGGGCCACCCCGCCAGCAGCATCTTCCAAAGT 345
 Db
 321 ACAAGGAGGAACATGAGTGGCTGTCTGGGGCCACCCCGCCAGCAGCATCTTCCAAAGT 380
 Qy
 346 CCACCGTGATCAACATCCACAGCGAGACTCCGTGCGCCGACCATGTCGTCGCTCCCTGT 405
 Db
 381 CCACCGTGATCAACATCCACAGCGAGACTCCGTGCGCCGACCATGTCGTCGCTCCCTGT 440
 Qy
 406 TCAACACCCCTCTTCTGAACTGGTGTCTCTGGGCTTCATAGCATTCGCGCTACTCCGTGA 465
 Db
 441 TCAACACCCCTCTTCTGAACTGGTGTCTCTGGGCTTCATAGCATTCGCGCTACTCCGTGA 500
 Qy
 466 AGTCTAGGACAGGAAGTGTGGCGAGTACCGGGGCCCGAGGCTATGCTCCACCG 525
 Db
 501 AGTCTAGGACAGGAAGTGTGGCGAGTACCGGGGCCCGAGGCTATGCTCCACCG 560
 Qy
 526 CCAAGTCCCTGAACATCTGGGCCCTGATCTGGGCATCTCATGACCATTTGGATTTCATCC 585
 Db
 561 CCAAGTCCCTGAACATCTGGGCCCTGATCTGGGCATCTCATGACCATTTGGATTTCATCC 620
 Qy
 586 TGTACTGGTATTCGGCTCTGTGACAGTCTACCATATTATTTACAGATAAATACAGGAAA 645
 Db
 621 TGTACTGGTATTCGGCTCTGTGACAGTCTACCATATTATTTACAGATAAATACAGGAAA 680
 Qy
 646 AACGGGTTACTAGTAGCGGCCATAGCTGCAACTTTGGCTCCACTCCAGTGCATGCTG 705
 Db
 681 AACGGGTTACTAGTAGCGGCCATAGCTGCAACTTTGGCTCCACTCCAGTGCATGCTG 740
 Qy
 706 GCCCTGCACGCTGGGGCTGTTCGCCCTTGGCTCCCTTGGCTCCCTTAGATACAGCAGTT 765
 Db
 741 GCCCTGCACGCTGGGGCTGTTCGCCCTTGGCTCCCTTGGCTCCCTTAGATACAGCAGTT 800
 Qy
 766 TATACCCACACACCTGTCTACAGTGTCTATCAATAAAGTG-ACGTGCTTGTGA 817
 Db
 801 TATACCCACACACCTGTCTACAGTGTCTATCAATAAAGTG-ACGTGCTTGTGA 853

RESULT 6
 AX409461
 LOCUS AX409461
 DEFINITION Sequence 2108 from Patent WO0229103.
 ACCESSION AX409461
 VERSION AX409461.1 GI:21442166
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
 1
 Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
 Gene expression profiles in liver cancer
 Patent: WO 0229103-A 2108 11-APR-2002;
 GENE LOGIC INC (US)
 LOCATION/Qualifiers
 1..853
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="EMBL/GenBank Accession No. J04164"
 ORIGIN

```

Query Match      62.8%; Score 696.2; DB 6; Length 853;
Best Local Similarity 99.4%; Pred. No. 2.3e-155;
Matches 709; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 106 AGCAGGAAATAGAACTTAAAGAGAAATACACACTTCTGAGAACTTGAACAGCAGGGGAA 165
DB 141 AGCAGGAAATAGAACTTAAAGAGAAATACACACTTCTGAGAACTTGAACAGCAGGGGAA 200

QY 166 AGGAGTCTCACTGAGCAGCGTCCAGCATCCGACACACAGCGCGCCCTTGCCTCCAG 225
DB 201 AGGAGTCTCACTGAGCAGCGTCCAGCATCCGACACACAGCGCGCCCTTGCCTCCAG 260

QY 226 CAGAAACACACACTTCTCAAACTTCACTCAACACTTCTTCCCCAAAGCCAGAGATGC 285
DB 261 CAGAAACACACACTTCTCAAACTTCACTCAACACTTCTTCCCCAAAGCCAGAGATGC 320

QY 286 ACNAGAGGAAACATGAGTGGCTGTCTGGGGGACACCCCGACGACCACTCTTCAAGGT 345
DB 321 ACNAGAGGAAACATGAGTGGCTGTCTGGGGGACACCCCGACGACCACTCTTCAAGGT 380

QY 346 CCACCGTGATCAACATCCACAGCAGACCTCCGTGCGCGACCATCTCGTCTGGTCCCTGT 405
DB 381 CCACCGTGATCAACATCCACAGCAGACCTCCGTGCGCGACCATCTCGTCTGGTCCCTGT 440

QY 406 TCAACACCCCTCTTCTTGAACCTGGTCTCTGGGGCTTCATAGCATTTCGCTTCCGCTGA 465
DB 441 TCAACACCCCTCTTCTTGAACCTGGTCTCTGGGGCTTCATAGCATTTCGCTTCCGCTGA 500

QY 466 AGTCTAGGACAGGAAGATGGTTGGCGAGCTGACCGGGGCCAGGCGCTATGCTCCACCG 525
DB 501 AGTCTAGGACAGGAAGATGGTTGGCGAGCTGACCGGGGCCAGGCGCTATGCTCCACCG 560

QY 526 CCAAGTCCCTGAAACATCTGGGCCCTGATTCTGGGCATCTCTATGACCATTTGGATTTCATCC 585
DB 561 CCAAGTCCCTGAAACATCTGGGCCCTGATTCTGGGCATCTCTATGACCATTTGGATTTCATCC 620

QY 586 TGTACTGGTATTTCGGCTCTGTGACAGCTTACCATATTATGTTTACAGATAATACAGAAA 645
DB 621 TGTACTGGTATTTCGGCTCTGTGACAGCTTACCATATTATGTTTACAGATAATACAGAAA 680

QY 646 AACGGGGTACTAGTACGCGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTG 705
DB 681 AACGGGGTACTAGTACGCGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTG 740

QY 706 GCCTTGACCGTGGGGCTGTGTCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 765
DB 741 GCCTTGACCGTGGGGCTGTGTCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 800

QY 766 TATACCCACACACCTGTCTACACTGACATTCAATAAAGTG-ACGTGCTTGTGA 817
DB 801 TATACCCACACACCTGTCTACAGTGTCAATTCAATAAAGTGACAGTGTGTGA 853

RESULT 7
LOCUS HUM927A
DEFINITION Human interferon-inducible protein 9-27 mRNA, complete cds.
ACCESSION J04164
VERSION J04164.1 GI:177801
KEYWORDS interferon-inducible protein.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 853)
AUTHORS Reid,L.E., Brannett,A.H., Gilbert,C.S., Porter,A.C., Gewert,D.R.,
Stark,G.N. and Kerr,I.M.
TITLE A single DNA response element can confer inducibility by both
alpha- and gamma-interferons
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 86 (3), 840-844 (1989)
MEDLINE 89128873
PubMed 2492664
COMMENT Original source text: Human lymphoid cell, cdna to mRNA, and DNA,

```

```

clones pHP[1,7,2,0,3,4].
Computer-readable sequence for [1] kindly provided by I.M.Kerr,
07-FEB-1989.
Bases 1-178 are of genomic origin. An intron is located between
positions 501 and 502 in the genomic DNA.
FEATURES
Location/Qualifiers
source 1..853
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
prim_transcript 177..852
prim_transcript 205..852
/note="9-27 mRNA"
/note="9-27 mRNA"
CDS 317..694
/note="interferon-inducible protein 9-27"
/codon_start=1
/protein_id="AAA35494.1"
/db_xref="GI:177802"
/translaton="MKKEHEVAVLGAPESTILPRSTVINIHSETSPDHHVWSLFNT
LFLNWCCLGFIATFAYSVKSRDRKMDVDTGAQAVASTAKCLNIWALILGILMTIGFIL
SLVFGSVTVVHIMLIQIEKRGY"
ORIGIN 156 bp upstream of AflII site.

Query Match      62.8%; Score 696.2; DB 9; Length 853;
Best Local Similarity 99.4%; Pred. No. 2.3e-155;
Matches 709; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 106 AGCAGGAAATAGAACTTAAAGAGAAATACACACTTCTGAGAACTTGAACAGCAGGGGAA 165
DB 141 AGCAGGAAATAGAACTTAAAGAGAAATACACACTTCTGAGAACTTGAACAGCAGGGGAA 200

QY 166 AGGAGTCTCACTGAGCAGCGTCCAGCATCCGACACACAGCGCGCCCTTGCCTCCAG 225
DB 201 AGGAGTCTCACTGAGCAGCGTCCAGCATCCGACACACAGCGCGCCCTTGCCTCCAG 260

QY 226 CAGAAACACACACTTCTCAAACTTCACTCAACACTTCTTCCCCAAAGCCAGAGATGC 285
DB 261 CAGAAACACACACTTCTCAAACTTCACTCAACACTTCTTCCCCAAAGCCAGAGATGC 320

QY 286 ACAAGGAGGAAACATGAGTGGCTGTGTGGGGCACACCCCGACGACCATCTTCCAAGGT 345
DB 321 ACAAGGAGGAAACATGAGTGGCTGTGTGGGGCACACCCCGACGACCATCTTCCAAGGT 380

QY 346 CCACCGTGATCAACATCCACAGCAGACCTCCGTGCGCGACCATCTCGTCTGGTCCCTGT 405
DB 381 CCACCGTGATCAACATCCACAGCAGACCTCCGTGCGCGACCATCTCGTCTGGTCCCTGT 440

QY 406 TCAACACCCCTCTTCTTGAACCTGGTCTCTGGGGCTTCATAGCATTTCGCTTCCGCTGA 465
DB 441 TCAACACCCCTCTTCTTGAACCTGGTCTCTGGGGCTTCATAGCATTTCGCTTCCGCTGA 500

QY 466 AGTCTAGGACAGGAAGATGGTTGGCGAGCTGACCGGGGCCAGGCGCTATGCTCCACCG 525
DB 501 AGTCTAGGACAGGAAGATGGTTGGCGAGCTGACCGGGGCCAGGCGCTATGCTCCACCG 560

QY 526 CCAAGTCCCTGAAACATCTGGGCCCTGATTCTGGGCATCTCTATGACCATTTGGATTTCATCC 585
DB 561 CCAAGTCCCTGAAACATCTGGGCCCTGATTCTGGGCATCTCTATGACCATTTGGATTTCATCC 620

QY 586 TGTACTGGTATTTCGGCTCTGTGACAGCTTACCATATTATGTTTACAGATAATACAGAAA 645
DB 621 TGTACTGGTATTTCGGCTCTGTGACAGCTTACCATATTATGTTTACAGATAATACAGAAA 680

QY 646 AACGGGGTACTAGTACGCGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTG 705
DB 681 AACGGGGTACTAGTACGCGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTG 740

QY 706 GCCTTGACCGTGGGGCTGTGTCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 765
DB 741 GCCTTGACCGTGGGGCTGTGTCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 800

QY 766 TATACCCACACACCTGTCTACACTGACATTCAATAAAGTG-ACGTGCTTGTGA 817

```

```
Db 801 TATACCCACACCTGCTACAGTGTCAATTAATAAGTGCACGTGTGGA 853
|||||
BC000897 693 bp mRNA linear PRI 12-NOV-2003
LOCUS BC000897
DEFINITION Homo sapiens interferon induced transmembrane protein 1 (9-27),
ACCESSION BC000897
KEYWORDS mRNA (CDNA clone MGC:5195 IMAGE:3464598), complete cds.
VERSION BC000897.1 GI:12654158
SOURCE MGC.
ORGANISM Homo sapiens (human)
REFERENCE 1 (bases 1 to 683)
AUTHORS Klausner, R.D., Collins, F.S., Wagner, L.H., Dergse, J.G.,
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 Strausberg, R.L., Feigold, E.A., Grouse, L.H., Dergse, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L.H., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Udgin, T.B., Toehiyuki, S.,
Carninci, P., Frange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalish, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
human and mouse cDNA sequences
Generation and initial analysis of more than 15,000 full-length
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16999-16903 (2002)
22388257
12477932
2 (bases 1 to 683)
Strausberg, R.
Direct Submission
Submitted (15-NOV-2000) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 4 Row: b Column: 24
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4504580.
Location/Qualifiers
1. .683
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:5195 IMAGE:3464598"
/tissue_type="Cervix, carcinoma"
/clone_lib="NIH_MGC_12"
FEATURES
source
```

```
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
1. .683
/genes="IFITM1"
/notes="synonyms: 9-27, IFI17, LEU13, CD225"
/db_xref="LocusID:8519"
/db_xref="MIM:604456"
131..508
/codon_start=1
/product="interferon induced transmembrane protein 1
(9-27)"
/protein_id="AAH00897.1"
/db_xref="GI:12654159"
/db_xref="LocusID:8519"
/translation="MKKEHEVAVIGAPPSTILPRSTVINHSETSVDPHVVSLPNT
LLNWCCLQFIAPYAVSKSDRMVGDVGTGAQYASTAKCLINAILILGILMTIGFIL
LLVFGSVTVYHMLQIIQKRGY"
188..433
/notes="CD225; Region: Interferon-induced transmembrane
protein. This family includes the human leukocyte antigen
CD225, which is an interferon inducible transmembrane
protein, and is associated with interferon induced cell
growth suppression"
/db_xref="CDD:pfam04505"
misc_feature
ORIGIN
Query Match 59.9%; Score 664.6; DB 9; Length 683;
Best Local Similarity 99.3%; Pred. No. 8.1e-148;
Matches 678; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
QY 152 AAACGACAGGGGAAAGAGGTCTCACTAGCAGCCGTCCTCCAGCATCCGAGACACACAGCGG 211
Db 1 AAACGACAGGGGAAAGAGGTCTCACTAGCAGCCGTCCTCCAGCATCCGAGACACACAGCGG 60
QY 212 CCCTTGCTCCAGCGAGAAACACACTTCTCAAACTTCACTCAACTTCTTCCCTCCCA 271
Db 61 CCCTTGCTCCAGCGAGAAACACACTTCTCAAACTTCACTCAACTTCTTCCCTCCCA 120
QY 272 AAGCCAGAGATGCACAAGGAGGAACATGAGTGGCTGTCTGGGGGACACCCCGAGCAC 331
Db 121 AAGCCAGAGATGCACAAGGAGGAACATGAGTGGCTGTCTGGGGGACACCCCGAGCAC 180
QY 332 CATCTTCCAAAGTCCACCGTGATCAACATCCACAGCGAGACCTCCGTCGCCGACCATGT 391
Db 181 CATCTTCCAAAGTCCACCGTGATCAACATCCACAGCGAGACCTCCGTCGCCGACCATGT 240
QY 392 CGTCTGCTCCGCTTCAACACCCCTCTTCTTGAAGTGGTGTCTGTGGGCTTCATAGCAT 451
Db 241 CGTCTGCTCCGCTTCAACACCCCTCTTCTTGAAGTGGTGTCTGTGGGCTTCATAGCAT 300
QY 452 CGCCTACTCCGTAAGTCTAGGAGACAGGAAGATGGTTGGGACGTCGACCGGGGCCAGGC 511
Db 301 CGCCTACTCCGTAAGTCTAGGAGACAGGAAGATGGTTGGGACGTCGACCGGGGCCAGGC 360
QY 512 CTATGCTCCACCGCAAGTCCCTGACATCTGGCCCTGATTCTGGGCATCCTCATGAC 571
Db 361 CTATGCTCCACCGCAAGTCCCTGACATCTGGCCCTGATTCTGGGCATCCTCATGAC 420
QY 572 CATTTGATTTCATCTCTTACTGTATTCTGGCTCTGTGACAGTCTACCATATATTATTACA 631
Db 421 CATTTGATTTCATCTCTTACTGTATTCTGGCTCTGTGACAGTCTACCATATATTATTACA 480
QY 632 GATATACAGGAAACCGGGTTACTAGTACCGCCCATAGCTGACACCTTTGCACCTCC 691
Db 481 GATATACAGGAAACCGGGTTACTAGTACCGCCCATAGCTGACACCTTTGCACCTCC 540
QY 692 ACTGTGCAATGTGGCCCTGCACGCTGGGGCTGTGTGCCCTCCCTTGGTCTTGCCCC 751
Db 541 ACTGTGCAATGTGGCCCTGCACGCTGGGGCTGTGTGCCCTCCCTTGGTCTTGCCCC 600
QY 752 TAGATACAGAGTTTATACCCACACACTCTCTACACTGACATTCATTAAGTG-ACGTG 810
Db 601 TAGATACAGAGTTTATACCCACACACTCTCTACAGTGTCTTCAATAAAGTGCACGTG 660
```

```

QY      811  CTTGTGAAAAAAGAAACAAATAAA 833
Db      661  CTTGTGAAAAAAGAAACAAATAAA 683

RESULT 9
HSR9A927      647 bp  mRNA  linear  PRI 17-FEB-1997
LOCUS      H.sapiens mRNA for interferon-induced 17kDa membrane protein.
ACCESSION  X84958
VERSION     9-27 gene; interferon-inducible protein; Leu-13 gene.
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1
AUTHORS    Deblandre, G.A., Marinx, O.P., Evans, S.S., Majja, S., Leo, O.,
           Caput, D., Huez, G.A. and Wathelot, M.G.
TITLE      Expression cloning of an interferon-inducible 17-kDa membrane
           protein implicated in the control of cell growth
JOURNAL    J. Biol. Chem. 270 (40), 23860-23866 (1995)
MEDLINE    96007544
PUBMED     7595564
REFERENCE   2 (bases 1 to 647)
AUTHORS    Deblandre, G.A.
TITLE      Direct Submission
JOURNAL    Submitted (22-FEB-1995) G.A. Deblandre, Universite Libre de
           Bruxelles, Departement de Biologie Moleculaire, 67, rue des
           Chevaux, B-1640 Rhode-St-Genese, BELGIUM
COMMENT    Sequence overlapping with that under the acc# J04164.
FEATURES   source
           1..647
           /organism="Homo sapiens"
           /mol_type="mRNA"
           /db_xref="taxon:9606"
           /cell_line="Hela"
           /clone_lib="ATCC CCL2"
           1..647
           /gene="9-27/Leu-13"
           74
           /gene="9-27/Leu-13"
           /note="J04164"
           /replace="a"
           111..488
           /gene="9-27/Leu-13"
           /codon_start=1
           /product="interferon-inducible protein"
           /protein_id="CAA59337.1"
           /db_xref="GI:1177476"
           /db_xref="GOA:P13164"
           /db_xref="SWISS-PROT:P13164"
           /translation="MKHEHEVAVLGAPPSTILPRSTVINIHSETSVDPHVWVSLENT
           LFNWCCLGFIAPYKSRDRKMGVDVTGAQYASTAKCLNIALILGILMTIGFIL
           LLYVFGSVTVYHIMLIQIKRGY"
           185
           /gene="9-27/Leu-13"
           /note="J04164"
           /replace="c"
           293
           /gene="9-27/Leu-13"
           /note="J04164"
           /replace="g"
           418
           /gene="9-27/Leu-13"
           /note="J04164"
           /replace="c"
           437
           /gene="9-27/Leu-13"
           /note="J04164"
           /replace="g"
           626..631
           /polyA_signal

```

```

ORIGIN
Query Match      56.3%; Score 624.4; DB 9; Length 647;
Best Local Similarity 98.9%; Pred. No. 3.2e-138;
Matches 639; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

/ gene="9-27/Leu-13"

QY      173  CTCACTGAGCACCGCTCCCGACATCGGACACACACAGCGCGCTTCGCTCCACCCAGAAA 232
Db      2    CTCACTGAGCACCGCTCCCGACATCGGACACACACAGCGCGCTTCGCTCCACCCAGAAA 61

QY      233  CCACACTTCTCAAAACCTTCACTCAACACTTCCTCCCAAGCCAGAGATGCACAGGA 292
Db      62  CCACACTTCTCATACCTTCACTCAACACTTCCTCCCAAGCCAGAGATGCACAGGA 121

QY      293  GGAACATGAGTGCGTGTGCTGGGGGACCCCGACACCATCTTCCAGGTTCCACCGT 352
Db      122  GGAACATGAGTGCGTGTGCTGGGGGACCCCGACACCATCTTCCAGGTTCCACCGT 181

QY      353  GATCAACATCCACAGCGAGAGCTCCGTCGCCGACCATGTCGTCTGCTCCCTGTTCAACAC 412
Db      182  GATTAACATCCACAGCGAGAGCTCCGTCGCCGACCATGTCGTCTGCTCCCTGTTCAACAC 241

QY      413  CCTCTTCTTGAACGTGCTGTCTGGGGCTTCATAGCATTCGCTACTCGTGAAGTCTAG 472
Db      242  CCTCTTCTTGAACGTGCTGTCTGGGGCTTCATAGCATTCGCTACTCGTGAAGTCTAG 301

QY      473  GGACAGAAAGATGTTGGCGACGTGACCGGGGCCAGGCGCTATGCTCCACCCCAAGTG 532
Db      302  GGACAGAAAGATGTTGGCGACGTGACCGGGGCCAGGCGCTATGCTCCACCCCAAGTG 361

QY      533  CTGGAACATCTGGGCCCTGATTCGGGCATCTCTATGACCATTGGATTATCCTGTTACT 592
Db      362  CTGGAACATCTGGGCCCTGATTCGGGCATCTCTATGACCATTGGATTATCCTGTTACT 421

QY      593  GSTATTGGCTCTGTGACAGTCTACCATATTATTTACAGATAATACAGAAAAACGGGG 652
Db      422  GSTATTGGCTCTGTATACAGTCTACCATATTATTTACAGATAATACAGAAAAACGGGG 481

QY      653  TTACTAGTAGCGGCCCATAGCTGCAACCTTTGCACTCCACGTGTGCAATGTCGCCCTGC 712
Db      482  TTACTAGTAGCGGCCCATAGCTGCAACCTTTGCACTCCACGTGTGCAATGTCGCCCTGC 541

QY      713  ACGTGGGGCTGTTGCCCTCCCTGCTGCTCCCTAGATACAGAGTTTATACCC 772
Db      542  ACGTGGGGCTGTTGCCCTCCCTGCTGCTCCCTAGATACAGAGTTTATACCC 601

QY      773  ACACACCTGTCTACACTGACATTCATTAATAAAGTG-ACGTGCTTGTGA 817
Db      602  ACACACCTGTCTACAGTGTCTTCAATAAAGTGACGTCGTTGTGA 647

RESULT 10
BT007173      378 bp  mRNA  linear  PRI 13-MAY-2003
LOCUS      Homo sapiens interferon induced transmembrane protein 1 (9-27)
DEFINITION  mRNA, complete cds.
ACCESSION  BT007173
VERSION     BT007173.1 GI:30593184
KEYWORDS   FLI_CDNA.
SOURCE     Homo sapiens (human)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 378)
AUTHORS    Kalnine, N., Chen, X., Rolfs, A., Halleck, A., Hines, L., Eisenstein, S.,
           Koundinya, M., Raphael, J., Moreira, D., Kelley, T., LaBaer, J., Lin, Y.,
           Phelan, M. and Farmer, A.
TITLE      Cloning of human full-length CDSs in BD Creator(TM) System Donor
           vector
JOURNAL    Unpublished
REFERENCE   2 (bases 1 to 378)
AUTHORS    Kalnine, N., Chen, X., Rolfs, A., Halleck, A., Hines, L., Eisenstein, S.,

```

Koundinya.M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y.,
Phelan,M. and Farmer,A.

TITLE JOURNAL

Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow
Circle, Palo Alto, CA 94303, USA

COMMENT
This CDS clone is a part of a collection of human full length
expression clones generated by BD Biosciences Clontech and the
Harvard Institute of Proteomics. Each CDS has been cloned in two
forms: with and without stop-codon (to allow fusion with C-terminal
tag). The CDS has been directionally cloned using BD In-Fusion(TM)
cloning system between the SalI and HindIII sites of the pDNR-DUAL
vector. Additional sequences in the clone: 'ACC' after SalI site
and before 'ATG' to provide Kozak consensus sequence; 'GG' after
last codon and before HindIII site to maintain reading frame.
Clone distribution: http://bioinfo.clontech.com/orfclones.

FEATURES

source

Location/Qualifiers
1..378
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GH00049X1.0"
/clone_lib="BD Creator(TM) CDS Library derived from MGC
collection"
/lab_host="DH5alpha T1 resistant"
/note="Vector: pDNR-Dual"
1..378
/product="interferon induced transmembrane protein 1
(9-27)"
/protein_id="AAP35837.1"
/db_xref="GI:30584731"
/translation="MHKEHEVAVILGAPSTILPRSTVINIHSETSVDPHVWMSLFNT
LFLNWCCLGFTAFAYSVKSRDRKMGVDVTGAQAYASTAKCLNIWALILGILMTIGFIL
LLVFGSVTVYHIMLQIIQEKRGY"

CDS

Query Match 34.1%; Score 378; DB 9; Length 378;
Best Local Similarity 100.0%; Pred. No. 2.1e-79;
Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 282 ATGCACAGGAGGAACATGAGGTGGCTGTCTGGGGGACCCCGCCAGCACCCTTCCA 341
DB 1 ATGCACAGGAGGAACATGAGGTGGCTGTCTGGGGGACCCCGCCAGCACCCTTCCA 60
QY 342 AGGTCCACCGTATCAACATCACAGCAGACCTCGTGGCCGACCATGTCGTGTCTCC 401
DB 61 AGGTCCACCGTATCAACATCACAGCAGACCTCGTGGCCGACCATGTCGTGTCTCC 120
QY 402 CTGTTTCAACACCTCTTCTTGAACCTGCTGTCTGGCTTTCATAGCATTCGCTACTCC 461
DB 121 CTGTTTCAACACCTCTTCTTGAACCTGCTGTCTGGCTTTCATAGCATTCGCTACTCC 180
QY 462 GTGAAGTCTAGGACAGGAGATGTTGGCGACGTGACCGGGGCCAGGCTATGCTCTCC 521
DB 181 GTGAAGTCTAGGACAGGAGATGTTGGCGACGTGACCGGGGCCAGGCTATGCTCTCC 240
QY 522 ACCGCCAGTGTCTGAACATCTGGGCTTCTGCGCATCTCATGACCATGGATTC 581
DB 241 ACCGCCAGTGTCTGAACATCTGGGCTTCTGCGCATCTCATGACCATGGATTC 300
QY 582 ATCTGTGTACTGGTATTCGGCTCTGTGACAGTCTACCATATTATGTTACAGATAATACAG 641
DB 301 ATCTGTGTACTGGTATTCGGCTCTGTGACAGTCTACCATATTATGTTACAGATAATACAG 360
QY 642 GAAACACGGGTACTAG 659
DB 361 GAAACACGGGTACTAG 378

ORIGIN

Query Match 34.1%; Score 378; DB 9; Length 378;
Best Local Similarity 100.0%; Pred. No. 2.1e-79;
Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 282 ATGCACAGGAGGAACATGAGGTGGCTGTCTGGGGGACCCCGCCAGCACCCTTCCA 341
DB 1 ATGCACAGGAGGAACATGAGGTGGCTGTCTGGGGGACCCCGCCAGCACCCTTCCA 60
QY 342 AGGTCCACCGTATCAACATCACAGCAGACCTCGTGGCCGACCATGTCGTGTCTCC 401
DB 61 AGGTCCACCGTATCAACATCACAGCAGACCTCGTGGCCGACCATGTCGTGTCTCC 120
QY 402 CTGTTTCAACACCTCTTCTTGAACCTGCTGTCTGGCTTTCATAGCATTCGCTACTCC 461
DB 121 CTGTTTCAACACCTCTTCTTGAACCTGCTGTCTGGCTTTCATAGCATTCGCTACTCC 180
QY 462 GTGAAGTCTAGGACAGGAGATGTTGGCGACGTGACCGGGGCCAGGCTATGCTCTCC 521
DB 181 GTGAAGTCTAGGACAGGAGATGTTGGCGACGTGACCGGGGCCAGGCTATGCTCTCC 240
QY 522 ACCGCCAGTGTCTGAACATCTGGGCTTCTGCGCATCTCATGACCATGGATTC 581
DB 241 ACCGCCAGTGTCTGAACATCTGGGCTTCTGCGCATCTCATGACCATGGATTC 300
QY 582 ATCTGTGTACTGGTATTCGGCTCTGTGACAGTCTACCATATTATGTTACAGATAATACAG 641
DB 301 ATCTGTGTACTGGTATTCGGCTCTGTGACAGTCTACCATATTATGTTACAGATAATACAG 360
QY 642 GAAACACGGGTACTAG 659
DB 361 GAAACACGGGTACTAG 378

RESULT 11

BT007946

LOCUS

DEFINITION Synthetic construct Homo sapiens mRNA linear SYN 13-MAY-2003

ACCESSION

BT007946.1 GI:30584730

VERSION

FLI CDNA.

KEYWORDS

SOURCE

ORGANISM

synthetic construct

artificial sequences

1 (bases 1 to 378)

REFERENCE

AUTHORS

Kalnine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S.,
Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y.,
Phelan,M. and Farmer,A.

TITLE

Cloning of human full-length CDSs in BD Creator(TM) System Donor
vector

UNPUBLISHED

2 (bases 1 to 378)

REFERENCE

AUTHORS

Kalnine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S.,
Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y.,
Phelan,M. and Farmer,A.

TITLE

Direct Submission

Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow
Circle, Palo Alto, CA 94303, USA

COMMENT

This CDS clone is a part of a collection of human full length
expression clones generated by BD Biosciences Clontech and the
Harvard Institute of Proteomics. Each CDS has been cloned in two
forms: with and without stop-codon (to allow fusion with C-terminal
tag). The CDS has been directionally cloned using BD In-Fusion(TM)
cloning system between the SalI and HindIII sites of the pDNR-DUAL
vector. Additional sequences in the clone: 'ACC' after SalI site
and before 'ATG' to provide Kozak consensus sequence; 'GG' after
last codon and before HindIII site to maintain reading frame.
Clone distribution: http://bioinfo.clontech.com/orfclones.

Location/Qualifiers

1..378
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"
/clone="GH00049X1.0"
/clone_lib="BD Creator(TM) CDS Library derived from MGC
collection"
/lab_host="DH5alpha T1 resistant"
/note="Vector: pDNR-Dual"
1..378
/note="Mutations: 377:Stop->Leu"
/codon_start=1
/transl_table=11
/product="Homo sapiens interferon induced transmembrane
protein 1 (9-27)"
/protein_id="AAP36618.1"
/db_xref="GI:30584731"
/translation="MHKEHEVAVILGAPSTILPRSTVINIHSETSVDPHVWMSLFNT
LFLNWCCLGFTAFAYSVKSRDRKMGVDVTGAQAYASTAKCLNIWALILGILMTIGFIL
LLVFGSVTVYHIMLQIIQEKRYL"

FEATURES

source

Location/Qualifiers

1..378
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"
/clone="GH00049X1.0"
/clone_lib="BD Creator(TM) CDS Library derived from MGC
collection"
/lab_host="DH5alpha T1 resistant"
/note="Vector: pDNR-Dual"
1..378
/note="Mutations: 377:Stop->Leu"
/codon_start=1
/transl_table=11
/product="Homo sapiens interferon induced transmembrane
protein 1 (9-27)"
/protein_id="AAP36618.1"
/db_xref="GI:30584731"
/translation="MHKEHEVAVILGAPSTILPRSTVINIHSETSVDPHVWMSLFNT
LFLNWCCLGFTAFAYSVKSRDRKMGVDVTGAQAYASTAKCLNIWALILGILMTIGFIL
LLVFGSVTVYHIMLQIIQEKRYL"

ORIGIN

Query Match 33.9%; Score 376.4; DB 12; Length 378;
Best Local Similarity 99.7%; Pred. No. 5.1e-79;
Matches 377; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 282 ATGCACAGGAGGAACATGAGGTGGCTGTCTGGGGGACCCCGCCAGCACCCTTCCA 341
DB 1 ATGCACAGGAGGAACATGAGGTGGCTGTCTGGGGGACCCCGCCAGCACCCTTCCA 60
QY 342 AGGTCCACCGTATCAACATCACAGCAGACCTCGTGGCCGACCATGTCGTGTCTCC 401
DB 61 AGGTCCACCGTATCAACATCACAGCAGACCTCGTGGCCGACCATGTCGTGTCTCC 120
QY 402 CTGTTTCAACACCTCTTCTTGAACCTGCTGTCTGGCTTTCATAGCATTCGCTACTCC 461
DB 121 CTGTTTCAACACCTCTTCTTGAACCTGCTGTCTGGCTTTCATAGCATTCGCTACTCC 180
QY 462 GTGAAGTCTAGGACAGGAGATGTTGGCGACGTGACCGGGGCCAGGCTATGCTCTCC 521
DB 181 GTGAAGTCTAGGACAGGAGATGTTGGCGACGTGACCGGGGCCAGGCTATGCTCTCC 240


```

/misc_feature      /note="assembly_fragment"
81235..89763
/misc_feature      /note="assembly_fragment"
89864..101889
/misc_feature      /note="assembly_fragment"
101990..112625
/misc_feature      /note="assembly_fragment"
112726..119802
/misc_feature      /note="assembly_fragment"
119903..125181
/misc_feature      /note="assembly_fragment"
125282..129184
/misc_feature      /note="assembly_fragment clone_end:T7 vector_side:right"
129285..137260
/misc_feature      /note="assembly_fragment"
137361..143542
/misc_feature      /note="assembly_fragment"
143643..146891
/misc_feature      /note="assembly_fragment"
146992..149568
/misc_feature      /note="assembly_fragment"
149669..152678
/misc_feature      /note="assembly_fragment"
152779..154135
/misc_feature      /note="assembly_fragment"
154236..155662
/misc_feature      /note="assembly_fragment"
155763..156813
/misc_feature      /note="assembly_fragment"

ORIGIN
Query Match      32.6%; Score 362; DB 2; Length 156819;
Best Local Similarity 100.0%; Pred. No. 1.1e-75;
Matches 362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 AGCAGGAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGAA 165
DB 78353 AGCAGGAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGAA 78294
QY 166 AGGAGGTCTACTGAGCAGCGTCCAGACATCCGGACACACAGCGGCCCTTCCTCCACG 225
DB 78293 AGGAGGTCTACTGAGCAGCGTCCAGACATCCGGACACACAGCGGCCCTTCCTCCACG 78234
QY 226 CAGAAACACACACTTCTCAAACTTCACTCAACACTTCTTCCCAAGCAGGAAGTGC 285
DB 78233 CAGAAACACACACTTCTCAAACTTCACTCAACACTTCTTCCCAAGCAGGAAGTGC 78174
QY 286 ACAGGAGGAACATGAGTGGTGTGCTGGGGGCACCCCGCAGCACCATCTTCCCAAGT 345
DB 78173 ACAGGAGGAACATGAGTGGTGTGCTGGGGGCACCCCGCAGCACCATCTTCCCAAGT 78114
QY 346 CCACCGTGATCAACATCCACAGCAGACCTCCGTGCCCGACCAATGTCGTGTCCTGT 405
DB 78113 CCACCGTGATCAACATCCACAGCAGACCTCCGTGCCCGACCAATGTCGTGTCCTGT 78054
QY 406 TCACACCTCTTCTGAACTGCTGTGCTGGGCTTCTAGCATTCGCTACTCCGTGA 465
DB 78053 TCACACCTCTTCTGAACTGCTGTGCTGGGCTTCTAGCATTCGCTACTCCGTGA 77994
QY 466 AG 467
DB 77993 AG 77992

RESULT 13
AP006286/c LOCUS
DEFINITION Homo sapiens chromosome 11 clone RP13-25N22 map 11p, WORKING DRAFT
SEQUENCE, 22 unordered pieces.
ACCESSION AP006286
VERSION AP006286.1 GI:29243346
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)

```

ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seog, P., Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Homo sapiens genomic DNA of 11p
JOURNAL Published Only in Database (2003)
REFERENCE 2 (bases 1 to 175416)
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seog, P., Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (24-MAR-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: hattori@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170)

Center: RIKEN Genomic Sciences Center (GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: HumDraft11
Center clone name: RP13-25N22
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 160074 bases at least Q40
Consensus quality: 167722 bases at least Q30
Consensus quality: 171662 bases at least Q20
Insert size: 173316; sum-of-contigs
Quality coverage: 7.92x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 22 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 33131 contig of 33131 bp in length
33232 54098 contig of 28667 bp in length
54199 72037 contig of 17839 bp in length
72138 83277 contig of 11140 bp in length
83378 98599 contig of 15222 bp in length
98700 105513 contig of 6814 bp in length
105614 116229 contig of 10616 bp in length
116330 121964 contig of 5635 bp in length
12065 127627 contig of 5563 bp in length
127728 131971 contig of 4244 bp in length
132072 137506 contig of 5435 bp in length
137607 143829 contig of 6223 bp in length
143930 148539 contig of 4610 bp in length
148640 153148 contig of 4509 bp in length
153249 157509 contig of 4261 bp in length
157610 160561 contig of 2952 bp in length
160662 164778 contig of 4117 bp in length
164879 168269 contig of 3391 bp in length
168370 170689 contig of 2320 bp in length
170790 172833 contig of 2044 bp in length
172934 174155 contig of 1222 bp in length
174255 175416 contig of 1161 bp in length.

* NOTE: This is a 'working draft' sequence. It currently consists of 22 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
1 33131: contig of 33131 bp in length

```
* 33132 33231: gap of 100 bp
* 33232 54098: contig of 20867 bp in length
* 54099 54198: gap of 100 bp
* 54199 72037: contig of 17839 bp in length
* 72038 83277: gap of 100 bp
* 83278 83377: contig of 11140 bp in length
* 83378 98599: gap of 100 bp
* 98600 98699: contig of 15222 bp in length
* 98700 105513: gap of 100 bp
* 105514 105613: contig of 6814 bp in length
* 105614 116229: gap of 100 bp
* 116230 116329: contig of 10616 bp in length
* 116330 121964: gap of 100 bp
* 121965 122064: contig of 5635 bp in length
* 122065 127627: gap of 100 bp
* 127628 127727: contig of 5563 bp in length
* 127728 131971: gap of 100 bp
* 131972 132071: contig of 4244 bp in length
* 132072 137506: gap of 100 bp
* 137507 137606: contig of 5435 bp in length
* 137607 143829: gap of 100 bp
* 143830 143929: contig of 6223 bp in length
* 143930 148539: gap of 100 bp
* 148540 148639: contig of 4610 bp in length
* 148640 153148: gap of 100 bp
* 153149 153248: contig of 4509 bp in length
* 153249 157509: gap of 100 bp
* 157510 157609: contig of 4261 bp in length
* 157610 160561: gap of 100 bp
* 160562 164778: gap of 100 bp
* 164779 164878: contig of 4117 bp in length
* 164879 168269: gap of 100 bp
* 168270 168369: contig of 3391 bp in length
* 168370 170689: gap of 100 bp
* 170690 170789: contig of 2320 bp in length
* 170790 172839: gap of 100 bp
* 172840 174155: contig of 2044 bp in length
* 174156 174256: gap of 100 bp
* 174257 175416: contig of 1222 bp in length
* 175417 175416: contig of 1161 bp in length.

FEATURES
    Location/Qualifiers
        1..175416
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="11"
            /map="11p"
            /clone="RP13-25N22"
        1..33131
            /note="assembly_fragment clone_end:SP6 vector_side:left"
        33232..54098
            /note="assembly_fragment"
        54199..72037
            /note="assembly_fragment"
        72138..83277
            /note="assembly_fragment"
        83378..98599
            /note="assembly_fragment"
        98700..105513
            /note="assembly_fragment clone_end:T7 vector_side:left"
        105614..116229
            /note="assembly_fragment"
        116330..121964
            /note="assembly_fragment"
        122065..127627
            /note="assembly_fragment"
        127728..131971
            /note="assembly_fragment"
        132072..137506
            /note="assembly_fragment"
        137607..143829
            /note="assembly_fragment"
```

```
misc_feature 143930..148539
              /note="assembly_fragment"
misc_feature 148640..153148
              /note="assembly_fragment"
misc_feature 153249..157509
              /note="assembly_fragment"
misc_feature 157610..160561
              /note="assembly_fragment"
misc_feature 160662..164778
              /note="assembly_fragment"
misc_feature 164879..168269
              /note="assembly_fragment"
misc_feature 168370..170689
              /note="assembly_fragment"
misc_feature 170790..172839
              /note="assembly_fragment"
misc_feature 172934..174155
              /note="assembly_fragment"
misc_feature 174256..175416
              /note="assembly_fragment"

ORIGIN
Query Match      32.6%; Score 362; DB 2; Length 175416;
Best Local Similarity 100.0%; Pred.No. 1.1e-75;
Matches 362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 AGCAGGAAATAGAACTTAAGAGAAATAACACACTTCTGAGAAACTGAAACGACAGGGGAA 165
    |||||
Db 76527 AGCAGGAAATAGAACTTAAGAGAAATAACACACTTCTGAGAAACTGAAACGACAGGGGAA 76468

QY 166 AGGAGGTCTCACTGAGCACCCTCCAGCATCCGAGACACACAGCGGCCCTTCGCTCCACG 225
    |||||
Db 76467 AGGAGGTCTCACTGAGCACCCTCCAGCATCCGAGACACACAGCGGCCCTTCGCTCCACG 76408

QY 226 CAGAAACACACACTTCTCAACCTTCACTCAACACTTCTTCCCAAGCCAGAGATGC 285
    |||||
Db 76407 CAGAAACACACACTTCTCAACCTTCACTCAACACTTCTTCCCAAGCCAGAGATGC 76348

QY 286 ACAAGGAGGAACATGAGGTGCTGTGTGGGGACCCGCCACACACACCTCTTCCAAGGT 345
    |||||
Db 76347 ACAAGGAGGAACATGAGGTGCTGTGTGGGGACCCGCCACACACCTCTTCCAAGGT 76288

QY 346 CCACCGTGATCAACATCCACAGCGAGCTCGTGCCCGACCATGCTGCTGCTGCTGCTGT 405
    |||||
Db 76287 CCACCGTGATCAACATCCACAGCGAGCTCGTGCCCGACCATGCTGCTGCTGCTGCTGT 76228

QY 406 TCAACACCTCTTCTTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 465
    |||||
Db 76227 TCAACACCTCTTCTTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 76168

QY 466 AG 467
    ||
Db 76167 AG 76166
```

RESULT 14

```
AC138230 187160 bp DNA linear PRI 22-MAY-2003
LOCUS Homo sapiens chromosome 11, clone RP13-317D12, complete sequence.
DEFINITION AC138230
ACCESSION AC138230
VERSION AC138230.5 GI:30984765
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 187160)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 11, clone RP13-317D12
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 187160)
AUTHORS Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
```

Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihoval,T., Mhova,T., Nguyen,C., Nicol,R., Norbu,C., Peterson,K.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Riese,C., Rogov,P., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talanas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
 Submitted (20-DEC-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 187160)

REFERENCE
 AUTHORS
 Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Maclean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihoval,T., Mhova,T., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Riese,C., Rogov,P., Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Vankataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
 Submitted (10-APR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 4 (bases 1 to 187160)

REFERENCE
 AUTHORS
 Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Maclean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihoval,T., Mhova,T., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Riese,C., Rogov,P., Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Vankataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
 Submitted (22-MAY-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On May 22, 2003 this sequence version replaced gi:29423936.
 All repeats were identified using RepeatMasker:
 Smit,A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L29013
 Center clone name: 317_D12

FEATURES	Location/Qualifiers
source	1..187160
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="11"
	/map="11"
	/clone="RP13-317D12"
	/clone_lib="RPC1-13 Human Female BAC"
repeat_region	135..200
	/rpt_family="L2"
repeat_region	1134..1310
	/rpt_family="MIR"
	/complement(1311..1614)
repeat_region	1615..1638
	/rpt_family="MIR"
repeat_region	1849..1938
	/rpt_family="L1ME"
repeat_region	complement(4297..4385)
	/rpt_family="L2"
repeat_region	4463..4561
	/rpt_family="FRAM"
repeat_region	complement(4573..4613)
	/rpt_family="MADE1"
repeat_region	4733..5040
	/rpt_family="AluY"
repeat_region	complement(5094..5383)
	/rpt_family="L2"
repeat_region	complement(5379..5505)
	/rpt_family="L2"
repeat_region	complement(5501..5627)
	/rpt_family="L2"
repeat_region	complement(5667..6096)
	/rpt_family="L2"
repeat_region	6542..6666
	/rpt_family="MER41E"
repeat_region	6699..6719
	/rpt_family="GC-rich"
repeat_region	6917..6996
	/rpt_family="C-rich"
repeat_region	7136..7329
	/rpt_family="C-rich"
repeat_region	15461..15483
	/rpt_family="TG)n"
repeat_region	18178..18413
	/rpt_family="L2"
repeat_region	complement(18543..18907)
	/rpt_family="MER4A"
repeat_region	complement(18913..19138)
	/rpt_family="AluSg/x"
repeat_region	complement(19139..19318)
	/rpt_family="MER4D"
repeat_region	complement(19319..19688)
	/rpt_family="MER4E1"
repeat_region	19689..19992
	/rpt_family="AluSp"
repeat_region	19994..20308
	/rpt_family="AluSg"
repeat_region	complement(20309..20630)
	/rpt_family="MER4E1"
repeat_region	20632..20763
	/rpt_family="AluSg/x"
repeat_region	complement(20785..20882)
	/rpt_family="MER4E"

Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachucka, A., Ramsaamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Strange-Thomann, N., Stojanovic, N., Stubbs, M., Tatalas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassilev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission

TITLE
JOURNAL
REFERENCE

Birren, B., Nuebaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Arachchi, H.M., Barna, N., Basien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Canarata, V., Chang, J., Chospel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Gagliardi, G., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J.P., Meneus, L., Minova, I., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramaeamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tsafaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M. Direct Submission
Submitted (08-JUL-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02144, USA
On Jul 8, 2003 this sequence version replaced gi:193978.
All repeats were identified using RepeatMasker:
Smith, A.F.P. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

TITLE	JOURNAL	COMMENT
1. The Role of the State in Economic Development	Journal of Economic Surveys	Excellent survey of the literature on the role of the state in economic development. The author provides a comprehensive overview of the theoretical and empirical literature, and offers a clear and concise summary of the findings.
2. The Impact of Globalization on the Environment	Environmental Economics and Policy	This article provides a detailed analysis of the impact of globalization on the environment. The author discusses the various channels through which globalization affects the environment, and provides a critical evaluation of the existing literature.
3. The Effect of Trade Liberalization on Income Inequality	Journal of International Trade and Development	This article examines the effect of trade liberalization on income inequality. The author uses a theoretical framework to analyze the impact of trade on income distribution, and provides empirical evidence to support the findings.
4. The Role of the Financial Sector in Economic Growth	Journal of Financial Economics	This article discusses the role of the financial sector in economic growth. The author provides a theoretical analysis of the relationship between the financial sector and economic growth, and offers policy recommendations based on the findings.
5. The Impact of Technology on the Labor Market	Journal of Labor Economics	This article examines the impact of technology on the labor market. The author discusses the various ways in which technology affects the labor market, and provides empirical evidence to support the findings.

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L28535
 Center clone name: 325 C 3

Only the first 143.8 kilobases of this clone are being submitted. The remainder overlaps accession number AC138230 [WICGR project L29013].

FEATURES	SOURCE
1. The first part of the document is a list of names and addresses, followed by a list of names and addresses.	1. The first part of the document is a list of names and addresses, followed by a list of names and addresses.
2. The second part of the document is a list of names and addresses, followed by a list of names and addresses.	2. The second part of the document is a list of names and addresses, followed by a list of names and addresses.
3. The third part of the document is a list of names and addresses, followed by a list of names and addresses.	3. The third part of the document is a list of names and addresses, followed by a list of names and addresses.
4. The fourth part of the document is a list of names and addresses, followed by a list of names and addresses.	4. The fourth part of the document is a list of names and addresses, followed by a list of names and addresses.
5. The fifth part of the document is a list of names and addresses, followed by a list of names and addresses.	5. The fifth part of the document is a list of names and addresses, followed by a list of names and addresses.
6. The sixth part of the document is a list of names and addresses, followed by a list of names and addresses.	6. The sixth part of the document is a list of names and addresses, followed by a list of names and addresses.
7. The seventh part of the document is a list of names and addresses, followed by a list of names and addresses.	7. The seventh part of the document is a list of names and addresses, followed by a list of names and addresses.
8. The eighth part of the document is a list of names and addresses, followed by a list of names and addresses.	8. The eighth part of the document is a list of names and addresses, followed by a list of names and addresses.
9. The ninth part of the document is a list of names and addresses, followed by a list of names and addresses.	9. The ninth part of the document is a list of names and addresses, followed by a list of names and addresses.
10. The tenth part of the document is a list of names and addresses, followed by a list of names and addresses.	10. The tenth part of the document is a list of names and addresses, followed by a list of names and addresses.

```
L290131. Location/Qualifiers
1. .143779
   /organism="Homo sapiens"
   /mol_type="genomic DNA"
   /db_xref="taxon:9606"
   /chromosome="11"
   /map="11"
   /clone="RP11-326C3"
   /clone_11b="BDC1-11 Human Male BAC"
```

```
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region
```

Query Match	32.5%	Score 360.4	DB 9	Length 143779
Best Local Similarity	99.7%	Pred. NO. 2.7e-75		
Matches 361	Conservative	0	Mismatches 1	Indels 0
				Gaps 0

Qy	106	AGCAGGAAATAGAAACTTAAAGAGAAATACACACTTCTGAGAAACTCGAAGCAGCAGGGGAA	165
Db	98539	AGCAGGAAATAGAAACTTAAAGAGAAATACACACTTCTGAGAAACTCGAAGCAGCAGGGGAA	98598
Qy	166	AGGAGGTCTCACTGAGCACCGTCCGAGATCCGGACACACAGCGGCCCTTCGCTCCACG	225
Db	98599	AGGAGGTCTCACTGAGCACCGTCCGAGATCCGGACACACAGCGGCCCTTCGCTCCACG	98658
Qy	226	CAGAAAACCACTTCGACACCTTCACACTTCCTCCCAAGAGCCAGAGATGC	285
Db	98659	CAGAAAACCACTTCGACACCTTCACACTTCCTCCCAAGAGCCAGAGATGC	98718

QY	286	ACAAGGAGGAACATGAGGTGGCTGTGCTGGGGGACACCCCGACACCATCTTCCAAGGT	345
Db	98719	ACAAGGAGGAACATGAGGTGGCTGTGCTGGGGGACACCCCGACACCATCTTCCAAGGT	98778
QY	346	CCACCGTGATCAACATCCACAGGAGACCTCCGTGCCCGACCATGTGCTGGTCCCTGT	405
Db	98779	CCACCGTGATCAACATCCACAGGAGACCTCCGTGCCCGACCATGTGCTGGTCCCTGT	98838
QY	406	TCAACACCCCTCTTCTTGAACCTGGTGTGCTGGGCTTCATAGCATTCGCCTACTCCGTGA	465
Db	98839	TCAACACCCCTCTTCTTGAACCTGGTGTGCTGGGCTTCATAGCATTCGCCTACTCCGTGA	98898
QY	466	AG 467	
Db	98899	AG 98900	

Search completed: March 11, 2004, 03:36:53
 Job time : 2828.74 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 10, 2004, 08:01:57 ; Search time 269.909 Seconds
(without alignments)
17454.974 Million cell updates/sec

Title: US-10-084-817-37
Perfect score: 1109
Sequence: 1 cctgcacaggagacactgg.....acanaattccaccctatccg 1109

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1051	94.8	1109	8	ACD42205 Human int
2	802	72.3	851	2	Aa41369 Human nor
3	703	63.4	1617	9	Ad225685 Human CDN
4	699.8	63.1	1095	6	Aa95008 Human DNA
5	696.2	62.8	842	6	Abk46192 cDNA enco
6	696.2	62.8	853	5	AbA83110 IFN-induc
7	696.2	62.8	853	6	Abk83770 Human cDN
8	696.2	62.8	853	6	Abk84497 Human ben
9	696.2	62.8	853	6	Abn95610 Gene #210
10	685.8	61.8	852	7	Abt31907 Human bre
11	624.4	56.3	647	6	Abq60782 Human 9-2
12	624.4	56.3	647	7	Abx10351 DNA enco
13	624.4	56.3	647	9	Abx18965 Human dis
14	562	50.7	643	6	Abq60669 Human col
15	553.4	49.9	651	9	Ad807174 Novel cod
16	552.8	49.8	674	6	Abq60529 Human col
17	541.4	48.8	619	6	Abq60439 Human col
18	518.8	46.8	579	6	Abq60650 Human col
19	511	46.1	640	6	Abq59690 Human col
20	481.6	43.4	527	6	Abq60421 Human col
21	480	43.3	581	6	Abq60275 Human col
22	464.4	41.9	676	6	Abq60312 Human col
23	454.8	41.0	701	6	Abq55533 Human ova

C	24	441.2	39.8	519	6	ABQ60383	Human col
	25	435	39.2	689	6	ABQ60308	Human col
	26	433	39.0	462	8	ACH32390	Human end
	27	427.6	38.6	434	8	ACH48424	Human leu
C	28	390.6	35.2	593	6	ABQ60272	Human col
	29	376.4	33.9	378	6	ABSS1531	Human cDN
	30	373.4	33.7	376	5	AAS77261	DNA enco
	31	373.4	33.7	376	5	AAS77263	DNA enco
	32	373.4	33.7	376	9	AD808931	Novel DNA
	33	364	32.8	384	5	AAS93007	DNA enco
	34	360.8	32.5	388	5	AAS75775	DNA enco
	35	355.8	32.1	598	6	ABQ60320	Human col
C	36	354	31.9	911	5	AAS75528	DNA enco
	37	353.2	31.8	507	4	AAS75528	Human col
	38	351	31.7	352	6	ABK39166	cDNA enco
	39	351	31.7	352	7	ACA11495	Human lun
	40	351	31.7	352	7	ACA11495	Human lun
	41	348.2	31.4	400	2	ACR02681	Lung canc
	42	340.8	30.7	622	9	AD22123	Human gen
	43	335.8	30.3	764	2	AD81113	Bovine le
C	44	335.8	30.3	764	3	AAV69033	DNA molec
C	45	335.8	30.3	764	6	AAC80995	Human bre
						AAS99841	Breast tu

ALIGNMENTS

RESULT 1
ID ACD42205 standard; cDNA; 1109 BP.
XX ACD42205;
AC ACD42205;
DT 05-SEP-2003 (first entry)
XX Human interferon inducible protein cDNA Incyte 3094768CB1.
DE Human; ss; DNA methylation; cancer; colon cancer.
XX Homo sapiens.
OS Homo sapiens.
XX US2003013099-A1.
XX 16-JAN-2003.
XX 07-MAR-2002; 2002US-00093766.
XX 19-MAR-2001; 2001US-0277380P.
XX (LASEK/) LASEK A K W.
XX (JONES/) JONES D A.
XX (KARPF/) KARPF A R.

Lasek AKW, Jones DA, Karpf AR;
WPI; 2003-503249/47.
P-PSDB; ABO25116.
XX New combination comprising cDNAs that are expressed in a disorder or process associated with DNA methylation, useful for diagnosing, staging, treating or monitoring treatment of cancer, e.g. colon cancer.
XX Claim 2; Page 22; 66pp; English.
XX The invention relates to a combination comprising cDNAs which are expressed in a disorder or process associated with DNA methylation. The combination and cDNAs are useful for diagnosing, staging, treating or monitoring treatment of cancer, e.g. colon cancer and for detecting changes in expression of genes encoding proteins that are associated with DNA methylation. The protein is useful for screening molecules or compounds to identify at least one ligand that binds to the protein and for producing an antibody. The present sequence represents a cDNA expressed in a disorder or process associated with DNA methylation

[illegible][illegible]

QY 61 CTCAGGAAAGAGCAAGATTTTCAGGAGAGAGCGCTGCGACAAAGTACAGCAGGAAATAGAAA 120
 DB 62 CTGAGGAAAGAGCAAGATTTTCAGGAGAGAGCGCTGCGACAAAGTACAGCAGGAAATAGAAA 121
 QY 121 CTTAAGAGAAATACACACTTCTGTAGAACTGAAACGACAGGAGGAGAGGTCTCACTGA 180
 DB 122 CTTAAGAGAAATACACACTTCTGTAGAACTGAAACGACAGGAGGAGAGGTCTCACTGA 180
 QY 181 GCACCGTCCAGCATCCGACACACACAGCGCGCTTGGCTCCACGACAGAAACCACTT 240
 DB 181 GCACCGTCCAGCATCCGACACACAGCGCGCTTGGCTCCACGACAGAAACCACTT 240
 QY 241 CTCACACCTTCACTCAACACTTCTTCCCAAGCCAGAAATGACAGGAGGAAACATG 300
 DB 241 CTCACACCTTCACTCAACACTTCTTCCCAAGCCAGAAATGACAGGAGGAAACATG 300
 QY 301 AGTGCGTGTGCTGGGGGACCCCGGACGACATCTTCCAAAGTCCACCGTATCAACA 360
 DB 301 AGTGCGTGTGCTGGGGGACCCCGGACGACATCTTCCAAAGTCCACCGTATCAACA 360
 QY 361 TCCACAGCAGAGCTCGTCCGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 DB 361 TCCACAGCAGAGCTCGTCCGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 QY 421 TGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 DB 421 TGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 QY 481 AGATGCTGCGAGCTGACCGGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 DB 481 AGATGCTGCGAGCTGACCGGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 QY 541 TCTGGGCGCTGATCTGGGCGATCTCATGACCATGATGATGATGATGATGATGATGATGAT 600
 DB 541 TCTGGGCGCTGATCTGGGCGATCTCATGACCATGATGATGATGATGATGATGATGATGAT 600
 QY 601 GCTCTGTGACAGTACAT 660
 DB 601 GCTCTGTGACAGTACAT 660
 QY 661 AGCGCGCCATAGCTGCAACCTTTGCACTCCACTGCTGCAATGCTGCGCTGCAACCTGCGG 720
 DB 661 AGCGCGCCATAGCTGCAACCTTTGCACTCCACTGCTGCAATGCTGCGCTGCAACCTGCGG 720
 QY 721 GCTGTTCCCGCTGCGCCCTTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 DB 721 GCTGTTCCCGCTGCGCCCTTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 QY 781 GTCTACACTGACATTCATTAAGTG-ACGTGCTTGTGAAAAAACAATAAACCCTGA 839
 DB 781 GTCTACAGTGTCAATTAAGTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 QY 840 GG 841
 DB 841 GG 842

RESULT 3
 ADE25685
 ID ADE25685 standard; cDNA; 1617 BP.
 XX
 AC ADE25685;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Human cDNA differentially expressed in foam cells #89.
 XX
 KW Human; ss; differential expression; foam cell; LPS; lipopolysaccharide;
 KW cardiovascular disease; atherosclerosis.
 XX
 OS Homo sapiens.
 XX
 FN US2003194721-A1.

XX 16-OCT-2003.
 PD 18-SEP-2002; 2002US-00247671.
 XX 19-SEP-2001; 2001US-0323784P.
 XX (INCY-) INCYTE GENOMICS INC.
 PA Mikita T, Shiffman D, Porter JG, Kaser MR;
 PI WPI; 2003-875398/81.
 DR
 XX
 PT Combination containing several polynucleotide that are differentially
 PT expressed in foam cells and complements of the polynucleotides, useful
 PT for diagnosing cardiovascular disease or atherosclerosis.
 XX
 PS Claim 1; SEQ ID NO 89; 37pp; English.
 XX
 CC The invention relates to a combination comprising several polynucleotides
 CC having any one of 127 sequences (S1) such as the sequence of human
 CC calmodulin gene, human mRNA for KIAA0930 protein, leukotriene A4
 CC hydrolase, human CGI-142 protein mRNA, human K+ channel beat 2 subunit
 CC mRNA, etc., and their complements. The cDNAs are differentially expressed
 CC in LPS (lipopolysaccharide)-treated foam cells. Also included are
 CC obtaining an extended or full length gene from a library of nucleic acid
 CC sequences, an expression vector containing the nucleic acids, a host cell
 CC containing the vector, a purified polypeptide appearing as ADE25750 and
 CC ADE25751, producing a protein by culturing the host cell, and a
 CC composition comprising a purified antibody that specifically binds to the
 CC proteins. The foam cell-expressed nucleic acids are useful for a high
 CC throughput detection of differential expression of one or more
 CC polynucleotides in a sample. The sample is from a subject with
 CC atherosclerosis and comparison with a standard defines early, mid or late
 CC stages of the disorder. The foam cell-expressed nucleic acids are useful
 CC for high throughput screening of a library of molecules or compounds to
 CC identify a ligand which binds a polynucleotide. The library is chosen
 CC from DNA molecules, peptides, proteins and RNA molecules. The protein is
 CC useful for a high throughput screening of library of molecules or
 CC compounds to identify at least one ligand which specifically binds a
 CC protein, for purifying a ligand from a sample for making an antibody. The
 CC foam cell-expressed nucleic acids are useful for diagnosing
 CC cardiovascular disorder. The foam cell-expressed nucleic acids are useful
 CC as elements on a microarray which can be used for detecting related
 CC polynucleotide in a sample, diagnosing cardiovascular disease,
 CC atherosclerosis. The present sequence represents a cDNA whose expression
 CC is upregulated in LPS treated foam cells.
 XX
 SQ Sequence 1617 BP; 348 A; 480 C; 445 G; 344 T; 0 U; 0 Other;
 Query Match 63.4%; Score 703; DB 9; Length 1617;
 Best Local Similarity 97.2%; Pred. No. 7.3e-182;
 Matches 726; Conservative 0; Mismatches 20; Indels 1; Gaps 1;
 QY 106 AGCAGGAAATAGAACTTAAGAGAAATACACTTCTGAGAACTGAAACGACAGGGGAA 165
 DB 379 AGCAGGAAATAGAACTTAAGAGAAATACACTTCTGAGAACTGAAACGACAGGGGAA 438
 QY 166 AGGAGGTTCTACTGAGCACCCTCCGACATCCGACACACAGCGGCCCTTCGCTCCACG 225
 DB 439 AGGAGGTTCTACTGAGCACCCTCCGACATCCGACACACAGCGGCCCTTCGCTCCACG 498
 QY 226 CAGAAAACACACTTCTCAAACTTCACTCAACACTTCTTCCCAAGCCAGAGATGC 285
 DB 499 CAGAAAACACACTTCTCAAACTTCACTCAACACTTCTTCCCAAGCCAGAGATGC 558
 QY 286 ACAGGAGGACATGAGGTGGTGGTGGGCGACCCCGACACCATCTTCCCAAGGT 345
 DB 559 ACAGGAGGACATGAGGTGGTGGTGGGCGACCCCGACACCATCTTCCCAAGGT 618
 QY 346 CCACCGTGATCAACATCCACAGCGAGACCTCGGTGCCCGACCATGTGCTGTGCTCCTGT 405
 DB 619 CCACCGTGATCAACATCCACAGCGAGACCTCGGTGCCCGACCATGTGCTGTGCTCCTGT 678

406	TCAACACCCCTCTCTTTGAACATCGTGTGTCTGGGCTTCATAGCATTCGCTACTCGGTGA	465
QY		
679	TCAACACCCCTCTCTTTGAACATCGTGTGTCTGGGCTTCATAGCATTCGCTACTCGGTGA	738
Db		
466	AGTCTAGGACAGGAAGATGGTTGGGACGTGACCGGGGCCAGGCGCTATGCCCTCCACCG	525
QY		
739	AGTCTAGGACAGGAAGATGGTTGGGACGTGACCGGGGCCAGGCGCTATGCCCTCCACCG	798
Db		
526	CCAAGTGCCTGGAACATCTGGGCCCTGATTCTGGGCATCTCATGACATTTGATGATTCA	585
QY		
799	CCAAGTGCCTGGAACATCTGGGCCCTGATTCTGGGCATCTCATGACATTTGATGATTCA	858
Db		
586	TGTTACTTGGTATTTCGGCTCTGTGCACAGTCTACCATATTATGTTACAGATAAATACAGAAA	645
QY		
859	TGTTACTTGGTATTTCGGCTCTGTGCACAGTCTACCATATTATGTTACAGATAAATACAGAAA	918
Db		
646	AACGGGGTTACTAGTAGCCGCCCATAGCCTGCAACCTTTGCATCTCACTGTGCAATGCTG	705
QY		
919	AACGGGGTTACTAGTAGCCGCCCATAGCCTGCAACCTTTGCATCTCACTGTGCAATGCTG	978
Db		
706	GCCTCTGCACGCTGGGGCTGTGTGCCCTTGCCCCCTTTGGTCTGTCGCCCTAGATACAGCAGTT	765
QY		
979	GCCTCTGCACGCTGGGGCTGTGTGCCCTTGCCCCCTTTGGTCTGTCGCCCTAGATACAGCAGTT	1038
Db		
766	TATACCCACACACCTGCTACATCGATTCATTAAGATG-AGTGTGCTTCTGTAAGAAAAA	824
QY		
1039	TATACCCACACACCTGCTACATCGATTCATTAAGATGCACTGTGTTGTGATGCT	1098
Db		
825	ACAAATAAAACCCGAGGGGGGGCCGG	851
QY		
1099	GTGACATTCATCTGGAGTGGGGGGCCCTG	1125
Db		

RESULT 4	
AAS95008	
ID	AAS95008 standard; DNA; 1095 BP.
XX	
AC	AAS95008;
XX	
DT	14-FEB-2002 (first entry)
XX	
DE	Human DNA sequence #263 expressed during foam cell differentiation.
XX	
KW	Human; foam cell differentiation; atherosclerosis; cerebral stroke;
KW	cardiovascular disorder; coronary artery disease; gene therapy; ds.
XX	
OS	Homo sapiens.
XX	
PN	WO200177389-A2.
XX	
PD	18-OCT-2001.
XX	
PF	04-APR-2001; 2001WO-US011128.
XX	
PR	05-APR-2000; 2000US-0195106P.
XX	
PA	(INCY-) INCYTE GENOMICS INC.
XX	
PI	Shiffman D, Somogyi R, Lawn R, Seilhamer JG, Porter GJ, Mikita T;
PI	Tai J;
XX	
DR	WPI; 2002-010925/01.
XX	
PT	Composition useful for diagnosis of conditions, disorders or diseases
PT	associated with atherosclerosis, comprises several polynucleotides that
PT	are differentially expressed in foam cell development.
XX	
PS	Claim 1; Page 303; 315pp; English.
XX	
CC	The present invention relates to the isolation of human polynucleotide
CC	sequences that are differentially expressed during foam cell
CC	differentiation. The polynucleotide sequences of the invention or a

CC	composition comprising these polynucleotides are useful as a high
CC	throughput method for detecting altered expression of one or more
CC	polynucleotides in a sample. The polynucleotides can be used in the
CC	diagnosis of disorders associated with cardiovascular disorders such as
CC	atherosclerosis, cerebral stroke, and cardiovascular disorders such as
CC	coronary artery disease. The polynucleotide sequences can also be used as
CC	PCR primers and probes. The polynucleotides of the invention are also
CC	useful in gene therapy. AAS94746-AAS95021 represent the human
CC	polynucleotide sequences of the invention which are differentially
CC	expressed during foam cell differentiation
XX	
SQ	Sequence 1095 BP; 267 A; 326 C; 268 G; 234 T; 0 U; 0 Other;
	Query Match 63.1%; Score 699.8; DB 6; Length 1095;
	Best Local Similarity 99.6%; Pred. No. 4,6e-181;
	Matches 712; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY	106 AGCAGGAATAGAACCTTAAGAGAAATACACACTTCTGAGAAAATCGAACCACAGGGGAA 165
Dd	381 AGCAGGAATAGAACCTTAAGAGAAATACACACTTCTGAGAAACTGAAACACACAGGGAA 440
QY	166 AGGAGGTCTCACTGAGCACCGTCCCAGCATCCGACACACAGCGGCCCTTCGCTCCAG 225
Dd	441 AGGAGGTCTCACTGAGCACCGTCCCAGCATCCGACACACAGCGGCCCTTCGCTCCAG 500
QY	226 CAGAAAACACACTTCTCAAACCTTCACATCAACACTTCTCCCCAAGCCAGAAATGC 285
Dd	501 CAGAAAACACACTTCTCAAACCTTCACATCAACACTTCTCCCCAAGCCAGAAATGC 560
QY	286 ACAAGGAGGAACATGAGGTGGTGCTGCGGGGCACCCCCAGCACCATCTTTCCAAGT 345
Dd	561 ACAAGGAGGAACATGAGGTGGTGCTGCGGGGCACCCCCAGCACCATCTTTCCAAGT 620
QY	346 CCACCGTGATCAACATCCACAGGAGACCTCGTGCCGACCATGTGCTGTCCTCCCTGT 405
Dd	621 CCACCGTGATCAACATCCACAGGAGACCTCGTGCCGACCATGTGCTGTCCTCCCTGT 680
QY	406 TCACACACCTCTCTTGAACTGTGTCTGTCTGGGCTTCATAGCATTCGCCCTACTCCGTGA 465
Dd	681 TCACACACCTCTCTTGAACTGTGTGTGTCTGGGCTTCATAGCATTCGCCCTACTCCGTGA 740
QY	466 AGTCTAGGACACAGGAAGATGGTTGGCGACGTGACCCGGGGCCAGGCCTATGCTCCACCG 525
Dd	741 AGTCTAGGACACAGGAAGATGGTTGGCGACGTGACCCGGGGCCAGGCCTATGCTCCACCG 800
QY	526 CCAAGTGCTGAACATCTGGGCCCTGATTCTGGGCATCCCTCATGACCATTTGGATTATCC 585
Dd	801 CCAAGTGCTGAACATCTGGGCCCTGATTCTGGGCATCCCTCATGACCATTTGGATTATCC 860
QY	586 TGTTACTTGATTTCCGCTCTGTGACAGTCTAACATATATTATGTTACAGATAATACAGGAAA 645
Dd	861 TGTTACTTGATTTCCGCTCTGTGACAGTCTAACATATATTATGTTACAGATAATACAGGAAA 920
QY	646 AACGGGGTTACTAGTAGGCCGCCCATAGCCTTGCAACCTTTGCACCTCCACTGTGCAATGCTG 705
Dd	921 AACGGGGTTACTAGTAGGCCGCCCATAGCCTTGCAACCTTTGCACCTCCACTGTGCAATGCTG 980
QY	706 GCCCTGCACGCTGGGGCTGTGGCCCTTGCCCTTGCTGCTGCCCTTAGATACAGCAGTT 765
Dd	981 GCCCTGCACGCTGGGGCTGTGGCCCTTGCCCTTGCTGCTGCCCTTAGATACAGCAGTT 1040
QY	766 TATACCCACACACCTGTCTACACTGCATTCATCAATAAAGTG-ACGTGCTTGTGAAA 819
Dd	1041 TATACCCACACACCTGTCTACAGTGCATTCATCAATAAAGTGACGCTGTGTGAAA 1095

RESULT 5	
ABK46192	
ID	ABK46192 standard; cDNA; 842 BP.
XX	
AC	ABK46192;
XX	
DT	05-JUN-2002 (first entry)

CC	expression of ovarian tumour marker genes (ABA83081-ABA83122, ABA83180,
CC	ABA83182 and ABA83184) or segments thereof (ABA83123-ABA83169, ABA83179,
CC	ABA83181 and ABA83183). The methods of the invention are useful for
CC	detecting an ovarian tumour in a patient, for identifying an individual
CC	at increased risk for developing ovarian cancer, in prognostic tests for
CC	assessing the relative severity of ovarian cancer, in tests for
CC	monitoring a patient in remission from ovarian cancer and in tests for
CC	monitoring disease status in a patient being treated for ovarian cancer.
CC	The methods can additionally be used to identify a particular tumour as
CC	being an ovarian tumour (i.e., an epithelial ovarian tumour selected from
CC	serous cystadenoma, borderline serous tumour, serous cystadenocarcinoma,
CC	mucinous cystadenoma, borderline mucinous tumour, mucinous
CC	cystadenocarcinoma, endometrioid carcinoma, undifferentiated carcinoma,
CC	clear cell adenocarcinoma, cystadenofibroma, adenofibroma and Brenner
CC	tumour. The ovarian tumour marker genes of the invention were identified
CC	using SAGE (serial analysis of gene expression) and were found to be
CC	overexpressed in a broad variety of ovarian epithelial tumour cells
CC	relative to normal ovarian epithelial cells. The marker genes are
CC	implicated in immune response pathways, in the regulation of cell
CC	proliferation and in protein folding, and many of these are membrane-
CC	localised or secreted. In addition to their use as diagnostic and
CC	prognostic markers, the ovarian tumour marker genes of their encoded
CC	proteins may be used as therapeutic targets for the treatment and
CC	prevention of ovarian cancer. Sequences ABA83081-ABA83122, ABA83180,
CC	ABA83182 and ABA83184 represent the ovarian tumour marker genes of the
CC	invention
XX	
SQ	Sequence 853 BP; 208 A; 268 C; 182 G; 195 T; 0 U; 0 Other;
	Query Match 62.8%; Score 696.2; DB 5; Length 853;
	Best Local Similarity 99.4%; Pred. No. 4e-180;
	Matches 709; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY	106 AGCAGGAAATAGAAACTTAAGAGAAATAACACACTTCTGAGAAACTGGAACGACGGGAA 165
DB	141 AGCAGGAAATAGAAACTTAAGAGAAATAACACACTTCTGAGAACTGGAACGACGGGAA 200
QY	166 AGGAGGTCTCACTGAGCACCGCTCCGAGCATCGGACACACAGCGGCCCTTCGTCTCCAG 225
DB	201 AGGAGGTCTCACTGAGCACCGCTCCGAGCATCGGACACACAGCGGCCCTTCGTCTCCAG 260
QY	226 CAGAAAACCACACTTCTCAAACCTTCACTCAACACTTCTCCCAAAGCCAGAAGATGC 285
DB	261 CAGAAAACCACACTTCTCAAACCTTCACTCAACACTTCTCCCAAAGCCAGAAGATGC 320
QY	286 ACAGGAGGACATGAGTGCTGTGCTGGGGGACCCCCAGACACCATCTTCCTCCAAGT 345
DB	321 ACAAGGAGGAAATAGTGTTGGTGTGTGCTGGGGGACCCCCAGACACCATCTTCCTCCAAGT 380
QY	346 CCACCGTGATCAACATCCACAGCGAGACCTCCGTGCCCGAACCATGTGCTCTGGTCCCTGT 405
DB	381 CCACCGTGATCAACATCCACAGCGAGACCTCCGTGCCCGAACCATGTGCTCTGGTCCCTGT 440
QY	406 TCACACACCTCTCTTGTAACCTGGTCTGCTGGGCTTTCATAGCATTCGGCTACTCCGTGA 465
DB	441 TCACACACCTCTCTCTTGTAACCTGGTCTGCTGGGCTTTCATAGCATTCGGCTACTCCGTGA 500
QY	466 AGTCTAGGGACAGGAAGATGTTGGCGACGTGACCGGGGCCCGAGGCCTATGCTCCACCG 525
DB	501 AGTCTAGGGACAGGAAGATGTTGGCGACGTGACCGGGGCCCGAGGCCTATGCTCCACCG 560
QY	526 CCAAGTGCGTGAACATCTGGGCCCTGATTCTGGGACCTCTCATGACCATTTGGAATTCATCC 585
DB	561 CCAAGTGCGTGAACATCTGGGCCCTGATTCTGGGCATCTCATGACCATTTGGAATTCATCC 620
QY	586 TGTTAATGTTATTCGGCTCTGTGACAGTCTACCATATTATGTTACAGATAATACAGGAAA 645
DB	621 TGTCATGTTATTCGGCTCTGTGACAGTCTACCATATTATGTTACAGATAATACAGGAAA 680
QY	646 AACGGGTTACTAGTAGCCGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTG 705
DB	681 AACGGGTTACTAGTAGCCGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTG 740

XX exposure of a subject to a pathogen or sterile inflammatory disease (e.g. psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal reperfusion injury, ARDS, adult respiratory distress syndrome, inflammatory bowel disease, Crohn's disease, ulcerative colitis, periodontal disease; also bacterial infection, viral infection, parasitic infection, protozoal infection, fungal infection and MS is useful for treating one of the above conditions. The present sequence represents a gene differentially expressed in granulocytes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 853 BP; 208 A; 268 C; 182 G; 195 T; 0 U; 0 Other;
Query Match 62.8%; Score 696.2; DB 6; Length 853;
Best Local Similarity 99.4%; Pred. No. 4e-180;
Matches 709; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

2Y 106 AGCAGGAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAA 165
Db 141 AGCAGGAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAA 200
Qy 166 AGGAGGTCTCACTGAGCACCCTCCAGCATCCGACACACAGCGGCCCTTCGCTCCAG 225
Db 201 AGGAGGTCTCACTGAGCACCCTCCAGCATCCGACACACAGCGGCCCTTCGCTCCAG 260
Qy 226 CAGAAACACACTTCTCAAACTTCACTCAACACTTCTCCCAAGCCAGAGATGC 285
Db 261 CAGAAACACACTTCTCAAACTTCACTCAACACTTCTCCCAAGCCAGAGATGC 320
Qy 286 ACAAGGAGAAATGAGGTGGTGTCTGGGGACCCCCAGCAACCATCTTCCAAAGT 345
Db 321 ACAAGGAGAAATGAGGTGGTGTCTGGGGACCCCCAGCAACCATCTTCCAAAGT 380
Qy 346 CCACGTGTCAACATCCACAGGACCTCCGTGCGGACCATGTCTGTGTCCTGT 405
Db 381 CCACGTGTCAACATCCACAGGACCTCCGTGCGGACCATGTCTGTGTCCTGT 440
Qy 406 TCAACACCTCTCTTGAACCTGTGTCTGGGCTTCATAGCATTCGCTTCCGTGA 465
Db 441 TCAACACCTCTCTTGAACCTGTGTCTGGGCTTCATAGCATTCGCTTCCGTGA 500
Qy 466 AGTCTAGGACAGGAATGGTGGGACGTGACCGGGGCCAGGCGCTATGCTCCACCG 525
Db 501 AGTCTAGGACAGGAATGGTGGGACGTGACCGGGGCCAGGCGCTATGCTCCACCG 560
Qy 526 CCAAGTGCTGACATCTGGGCTGATCTGGGCATCTCATGACCATTCGATTCATCC 585
Db 561 CCAAGTGCTGACATCTGGGCTGATCTGGGCATCTCATGACCATTCGATTCATCC 620
Qy 586 TGTACTGTGATTCGGCTCTGTGACAGTCTACCATATATATTACAGATATACAGGAAA 645
Db 621 TGTACTGTGATTCGGCTCTGTGACAGTCTACCATATATATTACAGATATACAGGAAA 680
Qy 646 ACAGGGTCTACTAGTACCGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTG 705
Db 681 ACAGGGTCTACTAGTACCGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTG 740
Qy 706 GCCTCTGACGCTGGGGCTGTGGCCCTTGCCTTGGCTTGGCTTGGCTTGGCTTGGCT 765
Db 741 GCCTCTGACGCTGGGGCTGTGGCCCTTGCCTTGGCTTGGCTTGGCTTGGCTTGGCT 800
Qy 766 TATACCCACACCTGTCTACACTGACATTCATTAATAAGTG-ACGTCCTTGTGA 817
Db 801 TATACCCACACCTGTCTACAGTGTCTATTCAATAAGTGACGCTTGTGA 853

RESULT 8
ABK64497
ID ABK64497 standard; DNA; 853 BP.
XX
AC ABK64497;

XX 18-JUN-2002 (first entry)
DT Human benign prostatic hyperplasia gene #392.
DE Human; benign prostatic hyperplasia; BPH; prostate cancer; gene; ds.
XX Homo sapiens.
OS WO200212440-A2.
FN 14-FEB-2002.
XX 07-AUG-2001; 2001WO-US024708.
XX 07-AUG-2000; 2000US-0223323P.
PR 05-JUN-2001; 2001US-00873319.
XX (GENE-) GENE LOGIC INC.
FA (NLSB) JAPAN TOBACCO INC.
XX Munger WE, Kulkarni P, Getzenberg RH, Waga I, Yamamoto J;
PI WPI; 2002-257476/30.
DR Identifying drugs for and diagnosing benign prostatic hyperplasia, by
XX detecting expression levels of one or more genes in prostate cells from
PT patient that are differentially regulated compared to normal prostate
PT cells.
XX Disclosure; Page 222; 444pp; English.
PS The invention relates to a method of diagnosing (I) the onset or
XX progression of benign prostatic hyperplasia (BPH), or screening (II) for
CC or identifying an agent that modulates the onset or progression of BPH.
CC The method is based on changes in gene expression in BPH tissue isolated
CC from patients exhibiting different clinical states of prostate
CC hyperplasia as compared to normal prostate tissue. (I) comprises
CC detecting the expression levels of one or more genes in prostate cells
CC from the subject that are differentially regulated compared to normal
CC prostate cells. (II) comprises preparing a first gene expression profile
CC of BPH cells or BPH-like cell population, exposing the cells to the
CC agent, and preparing a second gene expression profile of the agent exposed
CC cells, and comparing the first and second gene expression profiles. (I)
CC is useful for diagnosing the onset or progression of BPH. (II) is useful
CC for identifying an agent that modulates the onset or progression of BPH.
CC The methods are useful to present information identifying the expression
CC level in a tissue or cells, by comparing the expression level of genes
CC given in the specification in the tissue or cells to the level of
CC expression of gene in the database, and displaying the expression levels
CC of at least one gene in the tissue or cell sample compared to the
CC prostate cancer. ABK64106-ABK64860 represent human benign prostatic
CC hyperplasia gene sequences of the invention
XX Sequence 853 BP; 208 A; 268 C; 182 G; 195 T; 0 U; 0 Other;
SQ Query Match 62.8%; Score 696.2; DB 6; Length 853;
Best Local Similarity 99.4%; Pred. No. 4e-180;
Matches 709; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 106 AGCAGGAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAA 165
Db 141 AGCAGGAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAA 200
Qy 166 AGGAGGTCTCACTGAGCACCCTCCAGCATCCGACACACAGCGGCCCTTCGCTCCAG 225
Db 201 AGGAGGTCTCACTGAGCACCCTCCAGCATCCGACACACAGCGGCCCTTCGCTCCAG 260
Qy 226 CAGAAACACACTTCTCAAACTTCACTCAACACTTCTCCCAAGCCAGAGATGC 285
Db 261 CAGAAACACACTTCTCAAACTTCACTCAACACTTCTCCCAAGCCAGAGATGC 320

286 ACAAGAGGAAACATGAGTGGCTGTGTCTGGGGGACACCCAGCAGCACCATCTTCCAAAGT 345
 321 ACAGAGGAGGACATGAGTGGCTGTGTCTGGGGGACACCCAGCAGCACCATCTTCCAAAGT 380
 346 CCACCGTGATCAACATCCACAGGAGACCTCCGTCGGGACCATGTCGTGCTCCCTGT 405
 381 CCACCGTGATCAACATCCACAGGAGACCTCCGTCGGGACCATGTCGTGCTCCCTGT 440
 406 TCAACACCCCTCTTCTTGAACCTGGTGTCTGGGCTTCATAGCATTCGCCCTACTCCGTGA 465
 441 TCACAGACCCCTCTTCTTGAACCTGGTGTCTGGGCTTCATAGCATTCGCCCTACTCCGTGA 500
 466 AGCTTAGGGACAGAAAGATGGTTGGGACGTGACCGGGGCCAGCCCTATGCTCCACGG 525
 501 AGTCTAGGGACAGAAAGATGGTTGGGACGTGACCGGGGCCAGCCCTATGCTCCACGG 560
 526 CCAAGTGCCCTGAAACATCTGGGACCTGATCTGGGCTTCATAGCATTCGCCCTACTCCGTGA 585
 561 CCAAGTGCCCTGAAACATCTGGGACCTGATCTGGGCTTCATAGCATTCGCCCTACTCCGTGA 620
 586 TGTACTGTGATTCGGCTCTGTGACAGTCTACCATATTTATGTTACAGATAATACAGGAA 645
 621 TGTCACTGGTATTCGGCTCTGTGACAGTCTACCATATTTATGTTACAGATAATACAGGAA 680
 646 AACGGGGTTACTAGTACCGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTG 705
 681 AACGGGGTTACTAGTACCGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTG 740
 706 GCCTTGACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 765
 741 GCCTTGACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 800
 766 TATACCCACACACCTGCTACAGTGTCAATTAAGTG-ACGTGCTTGTGA 817
 801 TATACCCACACACCTGCTACAGTGTCAATTAAGTG-ACGTGCTTGTGA 853

RESULT 9
 ABN95610
 ID ABN95610 standard; DNA; 853 BP.
 XX
 AC ABN95610;
 XX
 DT 13-AUG-2002 (first entry)
 XX
 DE Gene #2108 used to diagnose liver cancer.
 XX
 KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotrophic;
 KW metastatic liver tumor; cytostatic; expression profile; disease state;
 KW disease progression; drug toxicity; drug efficacy; drug metabolism.
 XX
 OS Homo sapiens.
 XX
 PN W0200229103-A2.
 XX
 PD 11-APR-2002.
 XX
 PF 02-OCT-2001; 2001WO-US030589.
 XX
 PR 02-OCT-2000; 2000US-0237054P.
 XX
 PA (GENE-) GENE LOGIC INC.
 XX
 PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
 XX
 DR WPI; 2002-426119/45.
 XX
 PT Diagnosing and detecting the progression of liver cancer, hepatocellular
 PT carcinoma or metastatic liver tumor in a patient, involves detecting the
 PT level of expression of two or more genes in a liver tissue sample.
 XX
 PS Claim 1; SEQ ID NO 2108; 298pp; English.
 XX

CC The invention relates to a novel method for diagnosing and detecting the
 CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
 CC tumor in a patient, and differentiating metastatic liver cancer from
 CC hepatocellular carcinoma in a patient, involving detecting the level of
 CC expression of two or more genes represented in ABN93503-ABN97455 in a
 CC tissue sample. The method of the invention has hepatotropic, and
 CC cytostatic activity. The method is useful for diagnosing and detecting
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic
 CC liver carcinoma in a patient. The method is useful for identifying
 CC expression profiles which serve as useful diagnostic markers as well as
 CC markers that can be used to monitor disease states, disease progression,
 CC drug toxicity, drug efficacy and drug metabolism. Note: The sequence data
 CC for this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 853 BP; 208 A; 268 C; 182 G; 195 T; 0 U; 0 Other;
 Query Match 62.8%; Score 696.2; DB 6; Length 853;
 Best Local Similarity 99.4%; Pred. No. 4e-180;
 Matches 709; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
 QY 106 AGCAGGAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAACGACAGGGGAA 165
 DB 141 AGCAGGAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAACGACAGGGGAA 200
 QY 166 AGGAGGTCTCACTGAGCAGCAGTCCAGCATCCGACACACAGCGGCCCTTCGCTCCAG 225
 DB 201 AGGAGGTCTCACTGAGCAGCAGTCCAGCATCCGACACACAGCGGCCCTTCGCTCCAG 260
 QY 226 CAGAAACACACACTTCTCAACCTTCACTCAACCTTCTCCCAACGACGAGATGC 285
 DB 261 CAGAAACACACACTTCTCAACCTTCACTCAACCTTCTCCCAACGACGAGATGC 320
 QY 286 ACAAGGAGGAAACATGAGTGGCTGTGTGGGGGACACCCAGCAGCACCATCTTCCAAAGT 345
 DB 321 ACAAGGAGGAAACATGAGTGGCTGTGTGGGGGACACCCAGCAGCACCATCTTCCAAAGT 380
 QY 346 CCACCGTGATCAACATCCACAGCAGACCTCGTGCCTGGACCATGTCGTGCTCCCTGT 405
 DB 381 CCACCGTGATCAACATCCACAGCAGACCTCGTGCCTGGACCATGTCGTGCTCCCTGT 440
 QY 406 TCAACACCCCTCTTCTTGAACCTGGTGTCTGTGGGCTTCATAGCATTCGCCCTACTCCGTGA 465
 DB 441 TCAACACCCCTCTTCTTGAACCTGGTGTCTGTGGGCTTCATAGCATTCGCCCTACTCCGTGA 500
 QY 466 AGCTAGGGACAGAAAGATGGTTGGGACGTGACCGGGGCCAGCCCTATGCTCCACGG 525
 DB 501 AGTCTAGGGACAGAAAGATGGTTGGGACGTGACCGGGGCCAGCCCTATGCTCCACGG 560
 QY 526 CCAAGTGCCCTGAAACATCTGGGACCTGATCTGGGCTTCATAGCATTCGCCCTACTCC 585
 DB 561 CCAAGTGCCCTGAAACATCTGGGACCTGATCTGGGCTTCATAGCATTCGCCCTACTCC 620
 QY 586 TGTACTGTGATTCGGCTCTGTGACAGTCTACCATATTTATGTTACAGATAATACAGGAA 645
 DB 621 TGTCACTGGTATTCGGCTCTGTGACAGTCTACCATATTTATGTTACAGATAATACAGGAA 680
 QY 646 AACGGGGTTACTAGTACCGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTG 705
 DB 681 AACGGGGTTACTAGTACCGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTG 740
 QY 706 GCCTTGACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 765
 DB 741 GCCTTGACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 800
 QY 766 TATACCCACACACCTGCTACAGTGTCAATTAAGTG-ACGTGCTTGTGA 817
 DB 801 TATACCCACACACCTGCTACAGTGTCAATTAAGTG-ACGTGCTTGTGA 853
 RESULT 10
 ABT31907

CC	proteins. The DNA and protein sequences of the invention are useful as:			
CC	markers for tissues in which the corresponding protein is preferentially			
CC	expressed; as molecular weight markers on gels; as chromosome markers or			
CC	tags; to identify chromosomes or to map related gene positions; and to			
CC	compare with endogenous DNA sequences in patients to identify potential			
CC	genetic disorders. The present DNA sequence represents a gene of the			
CC	invention.			
XX				
SQ	Sequence 651 BP; 153 A; 205 C; 151 G; 142 T; 0 U; 0 Other;			
	Query Match 49.9%; Score 553.4; DB 9; Length 651;			
	Best Local Similarity 94.5%; Pred. No. 4.4e-141;			
	Matches 596; Conservative 0; Mismatches 31; Indels 4; Gaps 2;			
QY	157 ACAGGGGAAGAGGTCTC	ACTGAGCACCCTCCAGCATCCG	GACACACACAGCGGCCCTT	216
DB	3 ACGTGGGAAGAGGTCT	CACTGAGCACCCTCCAGCATCCG	GACACACAGCGGCCCTT	62
QY	217 CGCTCCAGCAGAAACCA	CACTTCAACCTTCACTCAAC	CTTCCCTCCCAAGCC	276
DB	63 CGCTCCAGCAGAAACCA	CACTTCAACCTTCACTCAAC	CTTCCCTCCCAAGCC	122
QY	277 AGAAGATGCACAGGAG	GAACATGAGTGGCTGTGTGG	GGGGCAACCCCGACACCATCC	336
DB	123 AGAAGATGCACAGGAG	GAACATGAGTGGCTGTGTGG	GGGGCAACCCCGACACCATCC	182
QY	337 TTCCAAGTCCACCGTGA	TCAACATCCACAGCAGACCT	CCGTGCCCGACCATGCTGCT	396
DB	183 TTCCAAGTCCACCGTGA	TCAACATCCACAGCAGACCT	CCGTGCCCGACCATGCTGACT	242
QY	397 GGTCCCTGTTCAACACC	CTTCTTGAACCTGGTGTCTG	TGGGCTTCATAGCATTGCGCT	456
DB	243 GGTCCCTGTTCAACACC	CTTCTTGAACCTGGTGTCTG	TGGGCTTCATAGCATTGCGCT	302
QY	457 ACTCCGTGAAGTCTAG	GGACAGGAAGTGTGGCGACG	TGACCGGGGCCAGGCCTATG	516
DB	303 ACTCCGTGAAGTCTAG	GGACAGGAAGTGTGGCGACG	TGACCGGGGCCAGGCCTATG	362
QY	517 CTTCCACCGCCCAAGT	CGCTGAACATCTGGGCCCTG	ATTCTGGGCATCCTCATGACCA	576
DB	363 CTTACACCCGCAAGT	CGCTGAACATCTGGGCCCTG	AAATCTGGGCATCCTCATGACCA	422
QY	577 GATTCACTCTGTTACT	GGTATTCTGGCTCTGTGACAG	TCTACCATATTATGTTACAGATA	636
DB	423 GATTCACTCTGTTACT	GGTATTCTGGCTCTGTGACAG	TCTACCATATTATGTTACAGATA	482
QY	637 TACAGAAACCGGGT	TACTAGTACCGCCCATAGC	TGCAACCTTGCACCTCCACTGT	696
DB	483 TACAGG---AACTGAG	TTACTAGTACCGCCCATAG	CTTGGAACCTTGCACCTCCACTGT	539
QY	697 GCAATGCTGCGCCCTG	CACGCTGGGGCTGTGGCCCT	TGCCCTTGGTCTCTGCCCTAGAT	756
DB	540 GCAATGCTGACCTGC	ACGCTGAAGCTTGTGCCCT	TGGTCTCTGCCCTAGAT	599
QY	757 ACAGCAGTTTATACCA	CACACACCTGTCTACA	787	
DB	600 ACAGCAG-TGATACCA	CACACACCTGTCTACA	629	

Search completed: March 10, 2004, 11:00:17
Job time : 273.909 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 10, 2004, 09:19:42 ; Search time 50.6434 Seconds
(without alignments)
12152.440 Million cell updates/sec

Title: US-10-084-817-37

Perfect score: 1109

Sequence: 1 cctgcacaggagacactgg.....acanaattccacccattccg 1109

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

- 1: /cgn2_6/ptodata/2/ina/5A COMB.seq:*
- 2: /cgn2_6/ptodata/2/ina/5B COMB.seq:*
- 3: /cgn2_6/ptodata/2/ina/6A COMB.seq:*
- 4: /cgn2_6/ptodata/2/ina/6B COMB.seq:*
- 5: /cgn2_6/ptodata/2/ina/PTUS COMB.seq:*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	351	31.7	352	4	US-09-702-705-1204 Sequence 1204, Ap
2	351	31.7	352	4	US-09-736-457-1204 Sequence 1204, Ap
3	351	31.7	352	4	US-09-614-124B-1204 Sequence 1204, Ap
4	351	31.7	352	4	US-09-671-325-1204 Sequence 1204, Ap
5	335.8	30.3	764	3	US-08-901-789A-283 Sequence 283, App
6	335.8	30.3	764	4	US-09-062-451-283 Sequence 283, App
7	335.8	30.3	764	4	US-09-289-198-283 Sequence 283, App
8	335.8	30.3	764	4	US-09-429-755-283 Sequence 283, App
9	281.8	25.4	808	4	US-09-023-655-1279 Sequence 1279, Ap
10	275.4	24.8	642	4	US-09-370-838-119 Sequence 119, App
11	175.6	15.8	207	4	US-09-702-705-998 Sequence 998, App
12	175.6	15.8	207	4	US-09-736-457-998 Sequence 998, App
13	175.6	15.8	207	4	US-09-614-124B-998 Sequence 998, App
14	175.6	15.8	207	4	US-09-671-325-998 Sequence 998, App
15	75	6.8	907	4	US-09-620-312D-652 Sequence 146, App
16	48	4.3	1285	4	US-09-149-476-146 Sequence 23, App1
17	46.4	4.2	4177	3	US-09-023-082A-23 Sequence 23, App1
18	46.4	4.2	4177	4	US-09-248-998-23 Sequence 476, App
19	43	3.9	1434	4	US-09-252-931A-476 Sequence 398, App
20	43	3.9	1995	4	US-08-232-931A-398 Sequence 64, App1
21	42.6	3.8	1517	4	US-09-904-615-64 Sequence 132, App
22	42.4	3.8	1088	4	US-09-148-545-132 Sequence 79, App1
23	42.4	3.8	1143	4	US-09-148-545-79 Sequence 27, App1
24	41.6	3.8	1002	4	US-09-257-179-27 Sequence 41, App1
25	41.4	3.7	2471	4	US-09-148-545-41 Sequence 34, App1
26	41	3.7	433	4	US-09-288-143-34 Sequence 8, App1
27	40.8	3.7	433	3	US-09-165-264-8

c	28	40.6	3.7	320	3	US-09-165-264-7	Sequence 7, Appli
c	29	40.6	3.7	320	3	US-09-165-264-13	Sequence 13, Appl
c	30	40.6	3.7	1080	4	US-09-149-476-222	Sequence 222, App
c	31	40.4	3.6	575	4	US-09-205-258-219	Sequence 219, App
c	32	40.4	3.6	1133	4	US-09-148-545-75	Sequence 75, Appl
c	33	40.4	3.6	1922	4	US-09-305-258-26	Sequence 26, Appl
c	34	40.2	3.6	426	4	US-09-252-991A-380	Sequence 380, App
c	35	40.2	3.6	1114	4	US-09-690-454-39	Sequence 39, Appl
c	36	39.6	3.6	1318	3	US-09-165-264-12	Sequence 12, Appl
c	37	39.6	3.6	1329	4	US-09-380-420C-1	Sequence 1, Appli
c	38	39.6	3.6	1264	4	US-09-899-642A-1	Sequence 32, Appl
c	39	39.4	3.6	7218	1	US-08-232-463-14	Sequence 14, Appl
c	40	39.4	3.6	1074	4	US-09-252-991A-15993	Sequence 15993, A
c	41	39.2	3.5	1287	4	US-09-252-991A-16510	Sequence 16510, A
c	42	39.2	3.5	1356	4	US-09-252-991A-16093	Sequence 16093, A
c	43	39.2	3.5	655	4	US-09-205-258-88	Sequence 88, Appl
c	44	39.2	3.5	667	4	US-09-288-143-36	Sequence 36, Appl
c	45	38.8	3.5	667	4	US-09-288-143-36	Sequence 36, Appl

ALIGNMENTS

RESULT 1

US-09-702-705-1204

; Sequence 1204, Application US/09702705

; Patent No. 6504010

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Lodes, Michael A.

; APPLICANT: Fanger, Gary

; APPLICANT: Vedvick, Tom

; APPLICANT: Carter, Darick

; APPLICANT: Retter, Marc

; APPLICANT: Mannion, Jane

; APPLICANT: Fan, Liqun

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER

; FILE REFERENCE: 210121.478C14

; CURRENT APPLICATION NUMBER: US/09/702,705

; CURRENT FILING DATE: 2000-10-30

; NUMBER OF SEQ ID NOS: 1833

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 1204

; LENGTH: 352

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-702-705-1204

Query Match 31.7% Score 351; DB 4; Length 352;

Best Local Similarity 100.0%; Pred. No. 1.3e-94;

Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 GGGGAAAGGAGGTCTCACTGAGCACCCTGCCAGATCCGGACACACAGCGGCCCTTCGC 219

Db 1 GGGGAAAGGAGGTCTCACTGAGCACCCTGCCAGATCCGGACACACAGCGGCCCTTCGC 60

QY 220 TCCACGCGAAGAACCCACACTTCTCAACCTTCACTCAACACTTCTTCTCCCAAGCCAGA 279

Db 61 TCCACGCGAAGAACCCACACTTCTCAACCTTCACTCAACACTTCTTCTCCCAAGCCAGA 120

QY 280 AGATGCACAAAGGAGGAACATGAGTGGCTGTGCTGGGGGACCCCGGACGACCATCTTC 339

Db 121 AGATGCACAAAGGAGGAACATGAGTGGCTGTGCTGGGGGACCCCGGACGACCATCTTC 180

QY 340 CAAGGTCCACCGTGATCAACATCCACAGCAGACCTCCGTGCGCCGACCATGTCTGGT 399

Db 181 CAAGGTCCACCGTGATCAACATCCACAGCAGACCTCCGTGCGCCGACCATGTCTGGT 240

QY 400 CCCTGTTTCAACACCCCTCTTCTTGAACCTGGTGTCTGGGCTTCATAGCATTGCGCTACT 459

Db 241 CCCTGTTTCAACACCCCTCTTCTTGAACCTGGTGTCTGGGCTTCATAGCATTGCGCTACT 300

QY 460 CCGTGAAGTCTAGGACAGGAAGATGGTTGGCGACGTGACCGGGGCCAGG 510
Db 301 CCGTGAAGTCTAGGACAGGAAGATGGTTGGCGACGTGACCGGGGCCAGG 351

RESULT 2

US-09-736-457-1204
; Sequence 1204, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fang, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736.457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1204
; LENGTH: 352
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-1204

Query Match 31.7%; Score 351; DB 4; Length 352;
Best Local Similarity 100.0%; Pred. No. 1.3e-94;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 GGGGAAGAGGAGTCTCACTGAGCACCGTCCAGCATCCGACACACACAGCGGCCCTTCGC 219
Db 1 GGGGAAGAGGAGTCTCACTGAGCACCGTCCAGCATCCGACACACACAGCGGCCCTTCGC 60
QY 220 TCCACGCGAGAAAACACACACTTCTCAAACTTCACTCAACACTTCTTCCCAAGCCAGA 279
Db 61 TCCACGCGAGAAAACACACACTTCTCAAACTTCACTCAACACTTCTTCCCAAGCCAGA 120
QY 280 AGATGCACAGGAGGAACATGAGTGGTGTGCTGGGGGACCCCGACACCATCTTC 339
Db 121 AGATGCACAGGAGGAACATGAGTGGTGTGCTGGGGGACCCCGACACCATCTTC 180
QY 340 CAAGGTCCACCGTGATCAACATCCACAGCAGACCTCCGTGCCCGACCATGTCTGGT 399
Db 181 CAAGGTCCACCGTGATCAACATCCACAGCAGACCTCCGTGCCCGACCATGTCTGGT 240
QY 400 CCTGTTCAACACCTCTTCTTGAACCTGTGCTGGGCTTCATAGCATTCGCCTACT 459
Db 241 CCTGTTCAACACCTCTTCTTGAACCTGTGCTGGGCTTCATAGCATTCGCCTACT 300
QY 460 CCGTGAAGTCTAGGACAGGAAGATGGTTGGCGACGTGACCGGGGCCAGG 510
Db 301 CCGTGAAGTCTAGGACAGGAAGATGGTTGGCGACGTGACCGGGGCCAGG 351

RESULT 3

US-09-614-124B-1204
; Sequence 1204, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fang, Gary
; APPLICANT: Vedvick, Tom

; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614.124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1204
; LENGTH: 352
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-614-124B-1204

Query Match 31.7%; Score 351; DB 4; Length 352;
Best Local Similarity 100.0%; Pred. No. 1.3e-94;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 GGGGAAGAGGAGTCTCACTGAGCACCGTCCAGCATCCGACACACACAGCGGCCCTTCGC 219
Db 1 GGGGAAGAGGAGTCTCACTGAGCACCGTCCAGCATCCGACACACACAGCGGCCCTTCGC 60
QY 220 TCCACGCGAGAAAACACACACTTCTCAAACTTCACTCAACACTTCTTCCCAAGCCAGA 279
Db 61 TCCACGCGAGAAAACACACACTTCTCAAACTTCACTCAACACTTCTTCCCAAGCCAGA 120
QY 280 AGATGCACAGGAGGAACATGAGTGGTGTGCTGGGGGACCCCGACACCATCTTC 339
Db 121 AGATGCACAGGAGGAACATGAGTGGTGTGCTGGGGGACCCCGACACCATCTTC 180
QY 340 CAAGGTCCACCGTGATCAACATCCACAGCAGACCTCCGTGCCCGACCATGTCTGGT 399
Db 181 CAAGGTCCACCGTGATCAACATCCACAGCAGACCTCCGTGCCCGACCATGTCTGGT 240
QY 400 CCTGTTCAACACCTCTTCTTGAACCTGTGCTGGGCTTCATAGCATTCGCCTACT 459
Db 241 CCTGTTCAACACCTCTTCTTGAACCTGTGCTGGGCTTCATAGCATTCGCCTACT 300
QY 460 CCGTGAAGTCTAGGACAGGAAGATGGTTGGCGACGTGACCGGGGCCAGG 510
Db 301 CCGTGAAGTCTAGGACAGGAAGATGGTTGGCGACGTGACCGGGGCCAGG 351

RESULT 4

US-09-671-325-1204
; Sequence 1204, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fang, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671.325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1204
; LENGTH: 352
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-671-325-1204

Query Match

31.7%; Score 351; DB 4; Length 352;

Best Local Similarity 100.0%; Pred. No. 1.3e-94;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 160 GGGGAAGAGGCTCTACTGAGACACCTCCAGCATCCGGACACACAGCGGCCCTTCGC 219
Db 1 GGGGAAGAGGCTCTACTGAGACACCTCCAGCATCCGGACACACAGCGGCCCTTCGC 60
QY 220 TCACGAGAAAACACACTTCTCAAACTTCACTCAAACTTCTTCCCAAGCCAGA 279
Db 61 TCACGAGAAAACACACTTCTCAAACTTCACTCAAACTTCTTCCCAAGCCAGA 120
QY 280 AGATGACAGAGGAGCAATGAGTGGCTGTCTGGGGGACCCCGCAGCACCATCTTC 339
Db 121 AGATGACAGAGGAGCAATGAGTGGCTGTCTGGGGGACCCCGCAGCACCATCTTC 180
QY 340 CAAGTCCACCGTATCAACATCCACAGCAGACCTCCGTGCGCCGACCATGCTGTGT 399
Db 181 CAAGTCCACCGTATCAACATCCACAGCAGACCTCCGTGCGCCGACCATGCTGTGT 240
QY 400 CCTGTTCACACCTCTTCTTGAACCTGTCTGTCTGGCTTCATAGCAATCGCTACT 459
Db 241 CCTGTTCACACCTCTTCTTGAACCTGTCTGTCTGGCTTCATAGCAATCGCTACT 300
QY 460 CCTGAGTCTAGGACAGGAGATGTTGCGAGCTGACCGGGGCCAGG 510
Db 301 CCTGAGTCTAGGACAGGAGATGTTGCGAGCTGACCGGGGCCAGG 351

RESULT 5
US-08-991-789A-283/c
; Sequence 283, Application US/08991789A
; Patent No. 6225054
; GENERAL INFORMATION:
; APPLICANT: Prudakis, Tony N.
; Smith, John M.
; Reed, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF BREAST CANCER
; NUMBER OF SEQUENCES: 292
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed IP Law Group
; STREET: 701 Fifth Avenue, Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/991,789A
; FILING DATE: 11-Dec-1997
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E. R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 210121.419C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 283:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 764 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 283:
US-08-991-789A-283

Query Match 30.3%; Score 335.8; DB 3; Length 764;
Best Local Similarity 92.9%; Pred. No. 6.6e-90;

Matches 341; Conservative 18; Mismatches 7; Indels 1; Gaps 1;
QY 463 TGAAGTCTAGGACAGGAAGATGTTGGCGACGTGACCGGGGCCCGAGCCTATGCTCCA 522
Db 367 TCAAGTCTAGGACAGGAGAGATGTTGGCGACGTGACCGGGGCCCGAGCCTATGCTCCA 308
QY 523 CCGCCAGTCTGAAACATCTGGGCCCTGATTCCTGGGCATCTCATGACCATTTGATTCA 582
Db 307 CCGCCAGTCTGAAATATGTTGGGCCCTGATTCCTGGGCATCTCATGACCATTTGATTCA 248
QY 583 TCTGTCTACTGTTATTCGGCTCTGTGACAGTCTACCATTTATTTAGTACAGATAATACAG 642
Db 247 TCTGTCTACTGTTATTCGGCTCTGTGACAGTCTACCATTTATTTAGTACAGATAATACAG 188
QY 643 AAAACGGGCTTACTAGTACGCGCCCATAGCTGCAACCTTTGGCACTTCCACTGTGCAATG 702
Db 187 AAAACGGGCTTACTAGTACGCGCCCATAGCTGCAACCTTTGGCACTTCCACTGTGCAATG 128
QY 703 CTGCGCCCTGACGCTGGGGCTGTTGCGCCCTGCGCCCTTGGTCCCTAGATACAGCA 762
Db 127 CTGCGCCCTGACGCTGGGGCTGTTGCGCCCTGCGCCCTTGGTCCCTAGATACAGCA 68
QY 763 GTTTATACCCACACACCTGCTACAGTCAATTAAGTG-ACGTGCTTTGTGAAAAA 821
Db 67 GTTTATACCCACACACCTGCTACAGTCAATTAAGTG-ACGTGCTTTGTGAAAAA 8
QY 822 AAAACAA 828
Db 7 AAAACAA 1

RESULT 6
US-09-062-451-283/c
; Sequence 283, Application US/09062451
; Patent No. 6344550
; GENERAL INFORMATION:
; APPLICANT: Prudakis, Tony N.
; Smith, John M.
; Reed, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF BREAST CANCER
; NUMBER OF SEQUENCES: 297
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/062,451
; FILING DATE: 04-APR-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Makl, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.419C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 283:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 764 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-062-451-283

Query Match 30.3%; Score 335.8; DB 4; Length 764;
 Best Local Similarity 92.9%; Pred. No. 6.6e-90;
 Matches 341; Conservative 18; Mismatches 7; Indels 1; Gaps 1;

QY 463 TGAAGTCTAGGACAGAGAGATGTTGGCGACGTGACCGGGGCCAGGCTATGCTCCA 522
 DB 367 TCAAGKGTAGGACAGAGAGAGATGTTGGCGACGTGACCGGGGCCAGGCTATGCTCCA 308
 QY 523 CGGCCAAGTGGCTGAACATCTGGGCGCTGATTCTGGGATCTCTCATGACCAITGGATTCA 582
 DB 307 CGGCCAAGTGGCTGAACATCTGGGCGCTGATTCTGGGATCTCTCATGACCAITGGATTCA 248
 QY 583 TCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCATATATTATGTTACAGATAATACAGG 642
 DB 247 TCSGTDAYTGGWATTCGGCTCTGTGACAGTCTACCATATATTATGTTACAGATAATACAGG 188
 QY 643 AAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAACCTTTGCACTCCACTGTGCAATG 702
 DB 187 AAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAACCTTTGCACTCCACTGTGCAATG 128
 QY 703 CTGCCCCGTGACCGTGGGGCTGTTGCCCTGCCCCCTGGTCTGCCCCCTAGATACAGCA 762
 DB 127 STGCCCCGTGACCGTGGGGCTGTTGCCCTGCCCCCTGGTCTGCCCCCTAGATACAGCA 68
 QY 763 GTTTATACCCACACACCTGTCTACACTGACATTCATTAATGAAGTG-ACGTGCTTGTGAAAA 821
 DB 67 GTTTATACCCACACACCTGTCTACACTGACATTCATTAATGAAGTGACGTGCTTGTGAAAA 8
 QY 822 AAAACAA 828
 DB 7 AAAAAAA 1

RESULT 7
 US-09-289-198-283/c
 ; Sequence 283, Application US/09289198
 ; Patent No. 6586570
 ; GENERAL INFORMATION:
 ; APPLICANT: Frudakis, Tony N.
 ; APPLICANT: Smith, John M.
 ; APPLICANT: Reed, Steven G.
 ; APPLICANT: Mishner, Lynda
 ; APPLICANT: Retter, Marc W.
 ; APPLICANT: Dillon, Davin C.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
 ; FILE OF INVENTION: TREATMENT AND DIAGNOSIS OF BREAST CANCER
 ; FILE REFERENCE: 210121.419C5
 ; CURRENT APPLICATION NUMBER: US/09/289,198
 ; CURRENT FILING DATE: 1999-04-09
 ; EARLIER APPLICATION NUMBER: US 09/062,451
 ; EARLIER FILING DATE: 1998-04-17
 ; EARLIER APPLICATION NUMBER: US 08/991,789
 ; EARLIER FILING DATE: 1997-12-11
 ; EARLIER APPLICATION NUMBER: US 08/838,762
 ; EARLIER FILING DATE: 1997-04-09
 ; EARLIER APPLICATION NUMBER: PCT/US97/00485
 ; EARLIER FILING DATE: 1997-01-10
 ; EARLIER APPLICATION NUMBER: US 08/700,014
 ; EARLIER FILING DATE: 1996-08-20
 ; EARLIER APPLICATION NUMBER: US 08/585,392
 ; EARLIER FILING DATE: 1996-01-01
 ; NUMBER OF SEQ ID NOS: 312
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 283
 ; LENGTH: 764
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(764)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-289-198-283

Query Match 30.3%; Score 335.8; DB 4; Length 764;
 Best Local Similarity 92.9%; Pred. No. 6.6e-90;
 Matches 341; Conservative 18; Mismatches 7; Indels 1; Gaps 1;

Matches 341; Conservative 18; Mismatches 7; Indels 1; Gaps 1;

QY 463 TGAAGTCTAGGACAGAGAGATGTTGGCGACGTGACCGGGGCCAGGCTATGCTCCA 522
 DB 367 TCAAGKGTAGGACAGAGAGAGATGTTGGCGACGTGACCGGGGCCAGGCTATGCTCCA 308
 QY 523 CGGCCAAGTGGCTGAACATCTGGGCGCTGATTCTGGGATCTCTCATGACCAITGGATTCA 582
 DB 307 CGGCCAAGTGGCTGAACATCTGGGCGCTGATTCTGGGATCTCTCATGACCAITGGATTCA 248
 QY 583 TCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCATATATTATGTTACAGATAATACAGG 642
 DB 247 TCSGTDAYTGGWATTCGGCTCTGTGACAGTCTACCATATATTATGTTACAGATAATACAGG 188
 QY 643 AAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAACCTTTGCACTCCACTGTGCAATG 702
 DB 187 AAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAACCTTTGCACTCCACTGTGCAATG 128
 QY 703 CTGCCCCGTGACCGTGGGGCTGTTGCCCTGCCCCCTGGTCTGCCCCCTAGATACAGCA 762
 DB 127 STGCCCCGTGACCGTGGGGCTGTTGCCCTGCCCCCTGGTCTGCCCCCTAGATACAGCA 68
 QY 763 GTTTATACCCACACACCTGTCTACACTGACATTCATTAATGAAGTG-ACGTGCTTGTGAAAA 821
 DB 67 GTTTATACCCACACACCTGTCTACACTGACATTCATTAATGAAGTGACGTGCTTGTGAAAA 8
 QY 822 AAAACAA 828
 DB 7 AAAAAAA 1

RESULT 8
 US-09-429-755-283/c
 ; Sequence 283, Application US/09429755A
 ; Patent No. 6656480
 ; GENERAL INFORMATION:
 ; APPLICANT: Frudakis, Tony N.
 ; APPLICANT: Smith, John M.
 ; APPLICANT: Reed, Steven G.
 ; APPLICANT: Mishner, Lynda
 ; APPLICANT: Retter, Marc W.
 ; APPLICANT: Dillon, Davin C.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
 ; FILE OF INVENTION: TREATMENT AND DIAGNOSIS OF BREAST CANCER
 ; FILE REFERENCE: 210121.419C6
 ; CURRENT APPLICATION NUMBER: US/09/429,755A
 ; CURRENT FILING DATE: 1999-10-28
 ; NUMBER OF SEQ ID NOS: 315
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 283
 ; LENGTH: 764
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(764)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-429-755-283

Query Match 30.3%; Score 335.8; DB 4; Length 764;
 Best Local Similarity 92.9%; Pred. No. 6.6e-90;
 Matches 341; Conservative 18; Mismatches 7; Indels 1; Gaps 1;

QY 463 TGAAGTCTAGGACAGAGAGATGTTGGCGACGTGACCGGGGCCAGGCTATGCTCCA 522
 DB 367 TCAAGKGTAGGACAGAGAGAGATGTTGGCGACGTGACCGGGGCCAGGCTATGCTCCA 308
 QY 523 CGGCCAAGTGGCTGAACATCTGGGCGCTGATTCTGGGATCTCTCATGACCAITGGATTCA 582
 DB 307 CGGCCAAGTGGCTGAACATCTGGGCGCTGATTCTGGGATCTCTCATGACCAITGGATTCA 248
 QY 583 TCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCATATATTATGTTACAGATAATACAGG 642

Db 247 TCSTGTDAYTGGWATTCGGCTCTGTGACAGTCTACCATATTATGTACAGATAATACAGG 188
 QY 643 AAAAAAGGGGTACTAGTAGCGCCCATAGCCTTGCAACCTTTGCACTCCACTGTGCAATG 702
 Db 187 AAAAAAGGGGTABTAGTAGCGCCCATAGCCTTGCAACCTTTGCACTCCACTGTGCAATG 128
 QY 703 CTGGCCGTGACCGTGGGGCTGTGGCCCTGCGCCCTTGGCTCTGCCCTAGATACAGCA 762
 Db 127 STGGCCCTGACCGTGGGGCTGTGGCCCTGCGCCCTTGGCTCTGCCCTAGATACAGCA 68
 QY 763 GTTTATACCCACACACCTGTCTACACTGACATTCAATAAAGTG-ACGTGCTTGTGAAAA 821
 Db 67 GTTTATACCCACACACCTGTCTACAGTGTCAATTCAATTAAGTGACAGTGTGCTTGGCGAAAA 8
 QY 822 AAAAAA 828
 Db 7 AAAAAA 1

RESULT 9
 US-09-023-655-1279
 ; Sequence 1279, Application US/09023655
 ; Patent No. 6607879
 ; GENERAL INFORMATION:
 ; APPLICANT: Cocks, Benjamin G.
 ; APPLICANT: Susan G. Stuart
 ; APPLICANT: Jeffrey J. Seilhamer
 ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
 ; TITLE OF INVENTION: EXPRESSION
 ; NUMBER OF SEQUENCES: 1508
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 ; STREET: 3174 PORTER DRIVE
 ; CITY: PALO ALTO
 ; STATE: CALIFORNIA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/023.655
 ; FILING DATE: HEREWITH
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Zeller, Karen J.
 ; REGISTRATION NUMBER: 37,071
 ; REFERENCE/DOCKET NUMBER: PA-0001 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (650) 855-0555
 ; TELEFAX: (650) 845-4166
 ; INFORMATION FOR SEQ ID NO: 1279:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 808 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: GENBANK
 ; CLONE: 9311374
 ; US-09-023-655-1279

Query Match 25.4%; Score 281.8; DB 4; Length 808;
 Best Local Similarity 80.2%; Pred. No. 7.6e-74;
 Matches 384; Conservative 0; Mismatches 82; Indels 13; Gaps 4;
 QY 128 GAAATACACACTTCTAGAAAACTGAAACGACAGGGGAAA-GGAGGTCTCACTGAG-CACC 185

Db 138 GAAAAAAGAACTGTGTGAAACCGAAACTACTGGGGAAGGGAGGCTCACTGAGTAACC 197
 QY 186 GTCCAGATCCGACACACAGCGGCGCTTCGTCTCCAGCGAGAAAAACACACTTCTCAA 245
 Db 198 ATCCAGTAACCG--ACCGCGCTGTGCTTCGTCTGGACACCATGAGTCACACTGTCCAA 255
 QY 246 ACCTTCACTCAACACTT-----CCTTCCCAAGCCAGAGATGCAAGAGGAGAA 296
 Db 256 ACCTTCTTCTCTCTGTCAACAGTGGCCAGCGCCCCCAACTATGAGATGCTCAAGAGGAG 315
 QY 297 CATGAGTGTGTGCTGGGGGACACCCCGACGACCATCTTCCAAAGTCCACCGTCAATC 356
 Db 316 CAGAGTGTGTGCTGGGGGGGCCCAACACCTGTCTCCCGACGTCACCGTGATC 375
 QY 357 AACATCCACAGGAGACCTTCGTGCGCGACCATGTCGTCTGCTGCTTCTTCAACACCTC 416
 Db 376 CACATCCGACGAGACCTTCGTGCGCGACCATGTCGTCTGCTGCTTCTTCAACACCTC 435
 QY 417 TTCTTGAACCTGTCTGCTGGGCTTCTATAGCATTCGCTACTCTCGTGAAGTCTAGGGAC 476
 Db 436 TTCAAGAACCCCTGCTGCTGGGCTTCTATAGCATTCGCTACTCGTGAAGTCTAGGGAC 495
 QY 477 AGGAAGATGTTGGCGACGTGACCGGGGCCGAGGCTATGCTCCACCGCCCAAGTGCCTG 536
 Db 496 AGGAAGATGTTGGCGACGTGACCGGGGCCGAGGCTATGCTCCACCGCCCAAGTGCCTG 555
 QY 537 AACATCTGGGCGCTGATTCTGGGCATCTCATGACCATGATTCATCTGTTACTTGGT 595
 Db 556 AACATCTGGGCGCTGATTCTGGGCATCTCATGACCATCTCTCATCTCATCTCCAGT 614

RESULT 10
 US-09-370-838-119
 ; Sequence 119, Application US/09370838
 ; Patent No. 8444425
 ; GENERAL INFORMATION:
 ; APPLICANT: Reed, Steven G.
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Mohamath, Roadoh
 ; APPLICANT: Secrist, Heather
 ; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
 ; TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE
 ; FILE REFERENCE: 210121.475C1
 ; CURRENT APPLICATION NUMBER: US/09/370.838
 ; CURRENT FILING DATE: 1999-08-09
 ; EARLIER APPLICATION NUMBER: US 09/285.323
 ; EARLIER FILING DATE: 1999-04-02
 ; NUMBER OF SEQ ID NOS: 289
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 119
 ; LENGTH: 642
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; US-09-370-838-119

Query Match 24.8%; Score 275.4; DB 4; Length 642;
 Best Local Similarity 70.0%; Pred. No. 5.4e-72;
 Matches 450; Conservative 0; Mismatches 146; Indels 47; Gaps 4;
 QY 202 ACCACAGCGGCCCTTCGTCTCCACGAGAGAAACACACTTCTCAAACTTCACTCAACACT 261
 Db 25 ACCGCGCTGTGCTTGGCTGGACACCATGATCACAATGTCCAAACCTTCTTCTCTCTG 84
 QY 262 T-----CCTTCCCCAAGCCAGAGATGCAAGAGGAGAACATGAGTGGTGGTGTGC 312
 Db 85 TCAACAGTGGCCAGCCCCCACTATGAGATGCTCAAGGAGGAGACGAGTGGTGTGC 144
 QY 313 TGGGGGACACCCCGCAGGACCATCTTCCAAGGTCCACCGTGTATCAATCCACAGCGAGA 372
 Db 145 TGGGGGCGCCCAACAACCTGCTCCCGGACGCTCCACCGTATCCACATCCGAGCAGA 204
 QY 373 CCTCGTGGCCGACCATGCTGCTGCTCTCTCTTCAACACCTCTTCTTGAATGGTGTCT 432

```

Db 205 CCTCGGCGCCGACATCGTCTGGTCCCTCTTCAACACCTCTTTCATGAACCCCTGCT 264
Qy 433 GTCTGGGCTTCATAGCATTGCGCTACTCCGTTGAAGTCTAGGACAGAGAGATGGTGGCG 492
Db 265 GCCTGGGCTTCATAGCATTGCGCTACTCCGTTGAAGTCTAGGACAGAGAGATGGTGGCG 324
Qy 493 ACGTGACGGGGCCGAGCCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGA 552
Db 325 ACGTGACGGGGCCGAGCCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGA 384
Qy 553 TTCTGGGCATCCTCATGACCATTTGATTCATCTGTTACTGTTATTCGGCTCTGTGACAG 612
Db 385 TTCTGGGCATCCTCATGACCATTTGCTCATCTGTCATCCAGTGTGATCTTCCAGSCCT 444
Qy 613 TCTACCATATTATGTTACAGATAATACAGAAACGGGGTACTAGTAGCCGCCCATAG 672
Db 445 ATGGATAGATCAGG-----AGGCATCACTGAGGCCAGGAGCTCTGCCCATGA 491
Qy 673 CCTCAACCTTTGCACTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTGGCCCT 732
Db 492 CCTGTATCCCAGTACTCACTTCCATCTCTCGCCCTG----- 530
Qy 733 GCQCCCTGTGCTGCGCCCTAGATACAGAGATTATATACCAACACACTGCTGTACACTGAC 792
Db 531 ---CCCCGGAGCGAGTCTGTATCAGCCCTTTATCTCTCACACGCTTTTCTACAATGCG 587
Qy 793 ATTCAATAAAGTG-ACGTGCTGTGCAAAACCAAAACAAATAAAA 834
Db 588 ATTCAATAAAGTGACGTTTCTGTGTAATAAATAAATAAATAA 630

```

RESULT 11

US-09-702-705-998
; Sequence 998, Application US/09702705
; Patent No. 6504810

GENERAL INFORMATION:

APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.478C14

CURRENT APPLICATION NUMBER: US/09/702,705
CURRENT FILING DATE: 2000-10-30

NUMBER OF SEQ ID NOS: 1833
SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 998
LENGTH: 207

TYPE: DNA
ORGANISM: Homo sapien

FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(207)

OTHER INFORMATION: n = A,T,C or G

US-09-702-705-998

Query Match 15.8%; Score 175.6; DB 4; Length 207;
Best Local Similarity 90.3%; Pred. No. 1.3e-42;
Matches 187; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

```

Qy 302 GGTGCTGTGCTGGGGGACACCCCGACACCTCTTCCAAAGTTCACCGTGATCAACAT 361
Db 1 GGTGCTGTGCTGGGGGCGGCGCCACAAACCTGTCTCCCGACGTCACCGTGATCCACAT 60

```

```

Qy 362 CCACAGCAGACCTCCGTCGCCCGACCATGTGGTCTGCTCCCTGTTCACACCCCTCTTCT 421

```

```

Db 61 NCGCAGCAGACCTCCGTCGCCCGACCATGTGCTGGTCCCTGTTCAACACCCCTCTTCT 120
Qy 422 GAATGCTGTGCTCTGGGCTTCATAGCAATTCGCTACTCGGTGAAGTCTAGGACAGGAA 481
Db 121 GAACCCCTGCTGCGCTGGGCTTCATAGCAATTCGCTACTCCGTGAAGTCTAGGACAGGAA 180
Qy 482 GATGTTGGCAGCTGACCGGGGCCCA 508
Db 181 GATGTTGGCAGCTGACCGGGGCCCA 207

```

RESULT 12

US-09-736-457-998
; Sequence 998, Application US/09736457
; Patent No. 6509448

GENERAL INFORMATION:

APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.478C15

CURRENT APPLICATION NUMBER: US/09/736,457
CURRENT FILING DATE: 2000-12-13

NUMBER OF SEQ ID NOS: 1884
SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 998
LENGTH: 207

TYPE: DNA
ORGANISM: Homo sapien

FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(207)

OTHER INFORMATION: n = A,T,C or G

US-09-736-457-998

Query Match 15.8%; Score 175.6; DB 4; Length 207;
Best Local Similarity 90.3%; Pred. No. 1.3e-42;
Matches 187; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

```

Qy 302 GGTGCTGTGCTGGGGGACACCCCGACACCTCTTCCAAAGTTCACCGTGATCAACAT 361
Db 1 GGTGCTGTGCTGGGGGCGGCGCCACAAACCTGTCTCCCGACGTCACCGTGATCCACAT 60

```

```

Qy 362 CCACAGCAGACCTCCGTCGCCCGACCATGTGGTCTGCTCCCTGTTCACACCCCTCTTCT 421
Db 61 NCGCAGCAGACCTCCGTCGCCCGACCATGTGGTCTGCTCCCTGTTCACACCCCTCTTCT 120

```

```

Qy 422 GAATGCTGTGCTCTGGGCTTCATAGCAATTCGCTACTCGGTGAAGTCTAGGACAGGAA 481
Db 121 GAACCCCTGCTGCGCTGGGCTTCATAGCAATTCGCTACTCCGTGAAGTCTAGGACAGGAA 180

```

```

Qy 482 GATGTTGGCAGCTGACCGGGGCCCA 508
Db 181 GATGTTGGCAGCTGACCGGGGCCCA 207

```

RESULT 13

US-09-614-124B-998
; Sequence 998, Application US/09614124B
; Patent No. 6630574

GENERAL INFORMATION:

APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary

Db 352 ACCGAGTGAACGACTATTACCTGGTCCATCTTCAACTTCGTCTACCTCAACTTCTGC 411
 Qy 432 TGTCTGGGCTTTCATAGCATTCGCTTACTCCGTGAAGTCTAGGACAGGAAGATGGTTGGC 491
 Db 412 TGCCTGGGCTTTCATCGCCTTGGCCTTACTCCCTCAAAGTCCGAGACAAGAGCTTCTCAAT 471
 Qy 492 GACGTGACCGGGGCCAGCCCTATGCCTCCACCGCCCAAGTGCCTGAACATCTGGGCCCTG 551
 Db 472 GACCTGAATGGAGCCGTGGAGATGCAAGACGGCCCGCTGTTCACATCACCAGTTCT 531
 Qy 552 ATTCTGGGCATCCTCATGACCATTGGATTTCATCTGTTACTG 593
 Db 532 GCCCTGGCAGCCTCCTGCATCATCCTCGTCTTCACTTCTG 573

Search completed: March 11, 2004, 14:41:57
 Job time : 52.6434 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2004, 03:37:21 ; Search time 258.418 Seconds
(without alignments)
15799.545 Million cell updates/sec

Title: US-10-084-817-37
Perfect score: 1109
Sequence: 1 cctgcaccaggagactgg.....acanaattccacctatccg 1109

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2432557 seqs, 1840798884 residues

Total number of hits satisfying chosen parameters: 4965114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq*
2:	/cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq*
3:	/cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq*
4:	/cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq*
5:	/cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq*
6:	/cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq*
7:	/cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq*
8:	/cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq*
9:	/cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq*
10:	/cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq*
11:	/cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq*
12:	/cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq*
13:	/cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq*
14:	/cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq*
15:	/cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq*
16:	/cgn2_6/ptodata/1/pubpna/US10C_NEW_PUB.seq*
17:	/cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq*
18:	/cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1051	94.8	1109	14	US-10-093-766-3
2	1051	94.8	1109	14	US-10-084-817-37
3	1051	94.8	1109	14	US-10-051-835-5
4	703	63.4	1617	14	US-10-247-671-89
5	699.8	63.1	1095	14	US-10-240-965-263
6	696.2	62.8	842	9	US-09-920-300A-1743
7	696.2	62.8	842	13	US-10-033-528-1743
8	696.2	62.8	842	14	US-10-099-926-1743
9	696.2	62.8	853	9	US-09-880-107-2108
10	696.2	62.8	853	10	US-09-960-706-826
11	696.2	62.8	853	10	US-09-873-319-392
12	696.2	62.8	853	14	US-10-171-581-195
13	696.6	61.9	1421	10	US-09-814-353-19857
14	685.8	61.8	852	14	US-10-176-847-27
15	624.4	56.3	647	14	US-10-301-822-80

16	624.4	56.3	647	15	US-10-133-937-25	Sequence 25, Appl
17	624.4	56.3	647	15	US-10-159-563-25	Sequence 25, Appl
18	465	41.9	468	12	US-10-085-783A-39396	Sequence 39396, A
19	465	41.9	468	15	US-10-242-535A-39396	Sequence 39396, A
20	456.4	41.2	470	12	US-10-085-783A-42090	Sequence 42090, A
21	456.4	41.2	470	15	US-10-242-535A-42090	Sequence 42090, A
22	454.8	41.0	701	15	US-10-264-049-1413	Sequence 1413, Ap
23	433	39.0	432	10	US-09-918-995-19602	Sequence 19602, A
24	427.6	38.6	434	10	US-09-918-995-35636	Sequence 35636, A
25	411	37.1	411	12	US-10-085-783A-32308	Sequence 32308, A
26	411	37.1	411	15	US-10-242-535A-32308	Sequence 32308, A
27	376.4	33.9	378	14	US-10-043-487-111	Sequence 111, App
28	353.2	31.8	507	14	US-10-106-698-2527	Sequence 2527, Ap
29	351	31.7	352	9	US-09-736-457-1204	Sequence 1204, Ap
30	351	31.7	352	9	US-09-902-941-1204	Sequence 1204, Ap
31	351	31.7	352	9	US-09-849-626-1204	Sequence 1204, Ap
32	351	31.7	352	14	US-10-017-754-1204	Sequence 1204, Ap
33	351	31.7	352	14	US-10-113-872-1204	Sequence 1204, Ap
34	350	31.6	350	9	US-09-796-692-5601	Sequence 5601, Ap
35	350	31.6	350	14	US-10-040-862-5601	Sequence 5601, Ap
36	350	31.6	350	15	US-10-057-475B-5601	Sequence 5601, Ap
37	350	31.6	350	15	US-10-154-884B-5601	Sequence 5601, Ap
38	340.8	30.7	622	14	US-10-299-497-2	Sequence 2, Appl
39	337.6	30.4	427	10	US-09-814-353-14867	Sequence 14867, A
40	335.8	30.3	764	9	US-09-810-936-383	Sequence 283, App
41	335.8	30.3	764	9	US-09-429-755-283	Sequence 283, App
42	335.8	30.3	764	9	US-09-924-400-283	Sequence 283, App
43	335.8	30.3	764	14	US-10-212-679-283	Sequence 283, App
44	332	29.9	341	9	US-09-796-692-3854	Sequence 3854, App
45	332	29.9	341	14	US-10-040-862-3854	Sequence 3854, Ap

ALIGNMENTS

RESULT 1
US-10-093-766-3
; Sequence 3, Application US/10093766
; Publication No. US20030013099A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.
; APPLICANT: Jones, David A.
; APPLICANT: Karpf, Adam R.
; TITLE OF INVENTION: GENES REGULATED BY DNA METHYLATION IN COLON TUMORS
; FILE REFERENCE: PA-0047 US
; CURRENT APPLICATION NUMBER: US/10/093,766
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 1109
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030013099A1 3094766CB1
; NAME/KEY: unsure
; LOCATION: 953, 963-996, 1002-1012, 1016-1018, 1022-1023, 1030, 1032-1034, 1053,
; LOCATION: 1061, 1093
; OTHER INFORMATION: a, t, c, g, or other
US-10-093-766-3

Query Match 94.8%; Score 1051; DB 14; Length 1109;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCTGCACGAGACACTGGGAGGTTTGTAGTCCCAAAACCCGACAGCAGGACTGCGAGC 60
Db 1 CCTGCACGAGACACTGGGAGGTTTGTAGTCCCAAAACCCGACAGCAGGACTGCGAGC 60
QY 61 CTGAGGAAAGCAAGGATTTTCAGGAGAGAGCGCTCGCAAGTGTAGCAGCAATAGAAA 120
Db 61 CTGAGGAAAGCAAGGATTTTCAGGAGAGAGCGCTCGCAAGTGTAGCAGCAATAGAAA 120

QY	121	CTTAAGAGAAAATACACACTTCTGAGAAA	CTGAAA	CGACAGGGGAAAGGAGGTCTCACTGA	180
DB	121	CTTAAGAGAAAATACACACTTCTGAGAAA	CTGAAA	CGACAGGGGAAAGGAGGTCTCACTGA	180
QY	181	GCACCGTCCAGCATCGGACACACACAGGGGCCCTT	CGCTCCACG	CAGAAAACACACACTT	240
DB	181	GCACCGTCCAGCATCGGACACACACAGGGGCCCTT	CGCTCCACG	CAGAAAACACACACTT	240
QY	241	CTCAAACTTCACTCAACACTTCTTCCGAAAGCCAGAA	GATGCA	CAAGGAGGAACATG	300
DB	241	CTCAAACTTCACTCAACACTTCTTCCGAAAGCCAGAA	GATGCA	CAAGGAGGAACATG	300
QY	301	AGGTGGCTGTGTGGGGGACACCCCGACGACCACT	CTCTTCCAA	GGTCCACGTCATCAACA	360
DB	301	AGGTGGCTGTGTGGGGGACACCCCGACGACCACT	CTCTTCCAA	GGTCCACGTCATCAACA	360
QY	361	TCACACGAGACCTCGGTGGCCGACCATGTGCTGTG	TCCCTG	TTCACACCCCTCTTCT	420
DB	361	TCACACGAGACCTCGGTGGCCGACCATGTGCTGTG	TCCCTG	TTCACACCCCTCTTCT	420
QY	421	TGAACCTGGTCTGTCTGGGCTTCATAGCAATTCGC	TCTAC	CTCCGTGAAAGTCTAGGGA	480
DB	421	TGAACCTGGTCTGTCTGGGCTTCATAGCAATTCGC	TCTAC	CTCCGTGAAAGTCTAGGGA	480
QY	481	AGATGTTGGCAGCTGACCGGGGCCAGGCTTATGCT	TCCACG	CGCAAGTGCCTGAACA	540
DB	481	AGATGTTGGCAGCTGACCGGGGCCAGGCTTATGCT	TCCACG	CGCAAGTGCCTGAACA	540
QY	541	TCGTGGCCCTGATTCCTGGGCATCCTCATGACCA	TTCGAT	TTCCTGTTACTGGTATTCG	600
DB	541	TCGTGGCCCTGATTCCTGGGCATCCTCATGACCA	TTCGAT	TTCCTGTTACTGGTATTCG	600
QY	601	GCCTCTGTGACAGTCTACCATATTTATGTATCAG	ATAATAC	AGGAAAAACGGGGTTACTAGT	660
DB	601	GCCTCTGTGACAGTCTACCATATTTATGTATCAG	ATAATAC	AGGAAAAACGGGGTTACTAGT	660
QY	661	AGCGGCCCATAGCCTTGCAACTTTTGCACTCCCA	CTGTG	CAATGTGCGCCTGCAAGCTGGG	720
DB	661	AGCGGCCCATAGCCTTGCAACTTTTGCACTCCCA	CTGTG	CAATGTGCGCCTGCAAGCTGGG	720
QY	721	GCCTGTTGCCCTCGCCCTTGGTCTGCTGCTTGT	GCCTT	TATACCCACACACTT	780
DB	721	GCCTGTTGCCCTCGCCCTTGGTCTGCTGCTTGT	GCCTT	TATACCCACACACTT	780
QY	781	GTCTACACTGACATTCAATAAGTGACGTGCTTGT	GTGAAAA	AAAAAACAATAAACCCGAG	840
DB	781	GTCTACACTGACATTCAATAAGTGACGTGCTTGT	GTGAAAA	AAAAAACAATAAACCCGAG	840
QY	841	GGGGGGCCCGGACCCCAATTCGCCCTAAGGGAG	GATATAC	ATTCCCGGGCGGTGTATAC	900
DB	841	GGGGGGCCCGGACCCCAATTCGCCCTAAGGGAG	GATATAC	ATTCCCGGGCGGTGTATAC	900
QY	901	ACGCTGGGATGGGACACCTTGGGTATCCATTA	ACGCT	TTCAGAGNCCCGGGG	960
DB	901	ACGCTGGGATGGGACACCTTGGGTATCCATTA	ACGCT	TTCAGAGNCCCGGGG	960
QY	961	ATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	NNNNNN	CCANNNNNNNNNNNNCCNNCC	1020
DB	961	ATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	NNNNNN	CCANNNNNNNNNNNNCCNNCC	1020
QY	1021	CNNCCACCCNNNNCCCCACTCTCCCCACAC	CACCCCA	CAACACACCCAGCCCTCA	1080
DB	1021	CNNCCACCCNNNNCCCCACTCTCCCCACAC	CACCCCA	CAACACACCCAGCCCTCA	1080
QY	1081	CCATCCCTCACAATAATTCACCCCTATCCG		1109	
DB	1081	CCATCCCTCACAATAATTCACCCCTATCCG		1109	

```

; Publication No. US20030119009A1
; GENERAL INFORMATION:
; APPLICANT: Susan Stuart
; APPLICANT: Jed G. Nuchtern
; APPLICANT: Sharon E. Pion
; APPLICANT: Jason M. Shohet
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
; FILE REFERENCE: PA-0046 US
; CURRENT APPLICATION NUMBER: US/10/084,817
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/270,784
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 365
; SOFTWARE: PERL Program
; SEQ ID NO 37
; LENGTH: 1109
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: incyte ID No. US20030119009A1 3094768C81
; NAME/KEY: unsure
; LOCATION: 953, 963-996, 1002-1012, 1016-1018, 1022-1023, 1030, 1032-1034, 1053,
; LOCATION: 1061, 1093
; OTHER INFORMATION: a, t, c, g, or other
; US-10-084-817-37

```

Query Match	94.8%;	Score 1051;	DB 14;	Length 1109;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1109; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	CCTGCACCAAGGACACTGGAGGTTTAGTCCCCAAACCGCGCACAGACAGGACTGTGACG	60	
DB	1	CCTGCACCAAGGACACTGGAGGTTTAGTCCCCAAACCGCGCACAGACAGGACTGTGACG	60	
QY	61	CTCAGGAAGAGCAAGGATTTACAGAGAGAGGCGCTGCCACAAGTGACACAGGAATAGAA	120	
DB	61	CTCAGGAAGAGCAAGGATTTACAGAGAGAGGCGCTGCCACAAGTGACAGGAATAGAA	120	
QY	121	CTTTAAGAGAAATACACACTTCTGAGAACTGTAAACGACAGGGGAAAGAGGTCTCACTGA	180	
DB	121	CTTTAAGAGAAATACACACTTCTGAGAACTGTAAACGACAGGGGAAAGAGGTCTCACTGA	180	
QY	181	GCACCGTCCAGCATTCGGACACACAGCGCGCGCTTCGCTCCAGCGAGAAAACACACATT	240	
DB	181	GCACCGTCCAGCATTCGGACACACAGCGCGCGCTTCGCTCCAGCGAGAAAACACACATT	240	
QY	241	CTCAAACTTCACTCAACACTTCTCTCCCAAGCCAGAAAGATGCACAAGGAGGAACATG	300	
DB	241	CTCAAACTTCACTCAACACTTCTCTCCCAAGCCAGAAAGATGCACAAGGAGGAACATG	300	
QY	301	AGTGTGCTGTGGGGGACACCCCGCAGCACATCTTCCAAGTCCACCGTGATCAACA	360	
DB	301	AGTGTGCTGTGTGGGGGACACCCCGCAGCACATCTTCCAAGTCCACCGTGATCAACA	360	
QY	361	TCCACAGCAGAGACTCGTGTCCCGACCATGTGCTCGTCCCTGTTCAACACCCCTCTTCT	420	
DB	361	TCCACAGCAGAGACTCGTGTCCCGACCATGTGCTCGTCCCTGTTCAACACCCCTCTTCT	420	
QY	421	TGAACCTGTGCTGTGTGGGTTTATAGCATTCGGCTACTCGTGAAGTCTAGGACAGGA	480	
DB	421	TGAACCTGTGCTGTGTGGGTTTATAGCATTCGGCTACTCGTGAAGTCTAGGACAGGA	480	
QY	481	AGATGGTTGGCGAGTGCACCGGGGCCACAGGCTATGCTCCACGCCAAAGTGCCTGAACA	540	
DB	481	AGATGGTTGGCGAGTGCACCGGGGCCACAGGCTATGCTCCACGCCAAAGTGCCTGAACA	540	
QY	541	TCGTGGGCCCTGTATCTGGGCATCTCATGACCATTGGATTTCATCTGTTACTGGTATTCG	600	
DB	541	TCGTGGGCCCTGTATCTGGGCATCTCATGACCATTGGATTTCATCTGTTACTGGTATTCG	600	
QY	601	GCTCTGTGACAGTCTACCATATTATGTTACAGATAATACAGGAAAAACGGGGTCTACTAGT	660	

601	GCTCTGTGACAGTCTACCATATATGTTACAGATAATACAGGAAAAACGGGGTTACTAGT	660
661	AGCGGCCATAGACCTGCAACCTTTGCACTCCACTGTGCAATGCTGCCCTGCAACGCTGGG	720
661	AGCGGCCATAGACCTGCAACCTTTGCACTCCACTGTGCAATGCTGCCCTGCAACGCTGGG	720
721	GCTGTTGCCCTGCCCCCTTGGTCTGCTGCCCTAGATACAGCAGTTTATACCCACACACCT	780
721	GCTGTTGCCCTGCCCCCTTGGTCTGCTGCCCTAGATACAGCAGTTTATACCCACACACCT	780
781	GTCTACACTGCATTCATTAAGTGCAGTGTGTGAAAAAACAATAAAACCCGAG	840
781	GTCTACACTGCATTCATTAAGTGCAGTGTGTGAAAAAACAATAAAACCCGAG	840
841	GGGGGGCGGACCCATTTCCGCCCTAAGGGAGGATATACATTCGCGGGCGGTGTATAC	900
841	GGGGGGCGGACCCATTTCCGCCCTAAGGGAGGATATACATTCGCGGGCGGTGTATAC	900
901	ACGCTGGGATGGGACACCTTGGGTATCCAAATTAACGCCCTGCATCTTCAGAGNCCCGGG	960
901	ACGCTGGGATGGGACACCTTGGGTATCCAAATTAACGCCCTGCATCTTCAGAGNCCCGGG	960
961	ATNN	1020
961	ATNN	1020
1021	CNNCCACCNCNNNNCCCACTCTCCCCACNCACCCCAACAAACACCCAGCCCTCA	1080
1021	CNNCCACCNCNNNNCCCACTCTCCCCACNCACCCCAACAAACACCCAGCCCTCA	1080
1081	CCATCCCTCCACANAATTCACCCCTATCCG	1109
1081	CCATCCCTCCACANAATTCACCCCTATCCG	1109

RESULT 3

US-10-051-835-5

Sequence 5, Application US/10051835

Publication No: US20030165864A1

GENERAL INFORMATION:

APPLICANT: Jones, David A.

APPLICANT: Laeak, Amy W.

TITLE OF INVENTION: GENES REGULATED BY DNA METHYLATION IN TUMOR CELLS

FILE REFERENCE: PA-0044 US

CURRENT APPLICATION NUMBER: US/10/051,835

CURRENT FILING DATE: 2002-01-16

NUMBER OF SEQ ID NOS: 25

SOFTWARE: PERL Program

SEQ ID NO 5

LENGTH: 1109

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: Inocyte ID No. US20030165864A1 3094768CB1

NAME/KEY: unsure

LOCATION: 953, 963-996, 1002-1012, 1016-1018, 1022-1023, 1030, 1032-1034, 1053,

LOCATION: 1061, 1093

OTHER INFORMATION: a, t, c, g, or other

US-10-051-835-5

	Query Match	94.8%	Score 1051	DB 14	Length 1109
	Best Local Similarity	100.0%	Pred. No. 0		
	Matches 1109	Conservative	Mismatches 0	Indels 0	Gaps 0
Qy	1	CCTGCACGAGGAGACACTGGGAGGTTTATGTCCTCCAAA	CCCGCACAGAGCAGGACTGTCAGC	60	
Db	1	CCTGCACCGAGGAGACACTGGGAGGTTTATGTCCTCCAAA	CCCGCACAGGAGGACTGTCAGC	60	
Qy	61	CTGAGGAAAGCAGCAAGGATTTTCAGGAGAGAGCCCTCGCA	CAAGTCAGTCAGCGAAATAGAAA	120	
Db	61	CTGAGGAAAGCAGCAAGGATTTTCAGGAGAGAGCCCTCGCA	CAAGTCAGTCAGCGAAATAGAAA	120	

RESULT 4
US-10-247-671-89
; Sequence 89, Application US/10247671
; Publication No. US20030194721A1

121	QY	121	CTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAAGAGGCTCTCACTGA	180
121	DB	121	CTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAAGAGGCTCTCACTGA	180
181	QY	181	GCACGGTCCCAGCATCCGGACACCAACGAGGGCCCTTCGCTCCACGACAGAAAACACACTT	240
181	DB	181	GCACGGTCCCAGCATCCGGACACCAACGAGGGCCCTTCGCTCCACGACAGAAAACACACTT	240
241	QY	241	CTCAAACTTCACTCAACACTTTCCTTCCCACAAAGCCAGAGATGCAACGAGGAAACATG	300
241	DB	241	CTCAAACTTCACTCAACACTTTCCTTCCCACAAAGCCAGAGATGCAACGAGGAAACATG	300
301	QY	301	AGGTGGCTGTGTGGGGGACACCCCGACGACCATCTTCCAAAGGTCACCGTGATCAACA	360
301	DB	301	AGGTGGCTGTGTGGGGGACACCCCGACGACCATCTTCCAAAGGTCACCGTGATCAACA	360
361	QY	361	TCACAGGAGACCTTCGTCGGGACCATGTGGTCGGTCCCTGTTCAACACCCCTCTTCT	420
361	DB	361	TCACAGGAGACCTTCGTCGGGACCATGTGGTCGGTCCCTGTTCAACACCCCTCTTCT	420
421	QY	421	TGAACTGTGTGTCTGGGCTTTCATAGCATTCGCCTACTCCGTGAAGTCTAGGGACAGGA	480
421	DB	421	TGAACTGTGTGTCTGGGCTTTCATAGCATTCGCCTACTCCGTGAAGTCTAGGGACAGGA	480
481	QY	481	AGATGGTTGGGACGTGACCGGGGCCAGCCCTATGCCTCCACCGCCAGTGCCTGAACA	540
481	DB	481	AGATGGTTGGGACGTGACCGGGGCCAGCCCTATGCCTCCACCGCCAGTGCCTGAACA	540
541	QY	541	TCTGGGCCCTGATTCCTGGGCATCCTCATGACCAATTGGATTCACTCTGTGTTACTGGTATTCG	600
541	DB	541	TCTGGGCCCTGATTCCTGGGCATCCTCATGACCAATTGGATTCACTCTGTGTTACTGGTATTCG	600
601	QY	601	GCTCTGTGACAGTCTACCAATATGTTACAGATAATACAGGAAACACGGGTACTAGT	660
601	DB	601	GCTCTGTGACAGTCTACCAATATGTTACAGATAATACAGGAAACACGGGTACTAGT	660
661	QY	661	AGCGGCCCATAGCCTGCAACCTTTGCACCTCCAATGCTGCAATGCTGGCCCTGCACGCTGGG	720
661	DB	661	AGCGGCCCATAGCCTGCAACCTTTGCACCTCCAATGCTGCAATGCTGGCCCTGCACGCTGGG	720
721	QY	721	GCTGTTGCCCTTGCCCTTGCTCTGCCCCATAGATACAGCAGTTTATACCCACACACCT	780
721	DB	721	GCTGTTGCCCTTGCCCTTGCTCTGCCCCATAGATACAGCAGTTTATACCCACACACCT	780
781	QY	781	GTCTACACTGACATTCAAATAAAGTGACGTGCTTGTGAAAAAATAAAATAAAACCCGAG	840
781	DB	781	GTCTACACTGACATTCAAATAAAGTGACGTGCTTGTGAAAAAATAAAATAAAACCCGAG	840
841	QY	841	GGGGGGCGGACCCCATTTCCCTTAAGGGGAGGATATACATTTCCGGGGGGGTGTTATAC	900
841	DB	841	GGGGGGCGGACCCCATTTCCCTTAAGGGGAGGATATACATTTCCGGGGGGGTGTTATAC	900
901	QY	901	ACGCTGGGATGGACACCTTGGGTATCCAATTAACGCTCATCCTTTCAGAGNCCCGGGG	960
901	DB	901	ACGCTGGGATGGACACCTTGGGTATCCAATTAACGCTCATCCTTTCAGAGNCCCGGGG	960
961	QY	961	ATNNCC	1020
961	DB	961	ATNNCC	1020
1021	QY	1021	CNNCCACCCNNCCCACTCTCCCCACACACCCCCANAAACAAACACCCAGCCCTCA	1080
1021	DB	1021	CNNCCACCCNNCCCACTCTCCCCACACACCCCCANAAACAAACACCCAGCCCTCA	1080
1081	QY	1081	CCATCCCTCACAAATTCACCTATCCG	1109
1081	DB	1081	CCATCCCTCACAAATTCACCTATCCG	1109

RESULT 4
US-10-24
; Sequen
; Public

GENERAL INFORMATION:
APPLICANT: Mikita, Thomas
APPLICANT: Shiffman, Dov
APPLICANT: Porter, Gordon, J.
APPLICANT: Kaser, Matthew R.
TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
FILE REFERENCE: PA-0050 US
CURRENT APPLICATION NUMBER: US/10/247,671
CURRENT FILING DATE: 2002-09-18
PRIOR APPLICATION NUMBER: 60/323,784
PRIOR FILING DATE: 2001-09-19
NUMBER OF SEQ ID NOS: 186
SOFTWARE: PERL Program
SEQ ID NO 89
LENGTH: 1617
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20030194721A1 1328372.14
US-10-247-671-89

Query Match 63.4%; Score 703; DB 14; Length 1617;
Best Local Similarity 97.2%; Pred. No. 1.9e-210;
Matches 726; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

QY 106 AGCAGGAAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAA 165
DB 379 AGCAGGAAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAA 438

QY 166 AGGAGGTCTACTGAGCAGCGTCCAGCATCCGACACACACAGCGGCGCTTGGCTCCACG 225
DB 439 AGGAGGTCTACTGAGCAGCGTCCAGCATCCGACACACACAGCGGCGCTTGGCTCCACG 498

QY 226 CAGAAACACACACTTCAAACTTCACTCAACACTTCTTCCCAAGCCAGCAATCCCTTCCAAAGGT 295
DB 499 CAGAAACACACACTTCAAACTTCACTCAACACTTCTTCCCAAGCCAGCAATCCCTTCCAAAGGT 558

QY 286 ACAAGGAGAAATAGAGGTGGTGTCTGGGGGACACCCCGGACACCATCTTCCAAAGGT 345
DB 559 ACAAGGAGAAATAGAGGTGGTGTCTGGGGGACACCCCGGACACCATCTTCCAAAGGT 618

QY 346 CCACCGTGATCAACATCCACAGCAGACCTCCGTGCGCCGACCATCTGCTGGTCCCTGT 405
DB 619 CCACCGTGATCAACATCCACAGCAGACCTCCGTGCGCCGACCATCTGCTGGTCCCTGT 678

QY 406 TCAACACCTCTTCTTGAACCTGGTGTCTGGGCTTCATAGCATTCGCTACTCCGTGA 465
DB 679 TCAACACCTCTTCTTGAACCTGGTGTCTGGGCTTCATAGCATTCGCTACTCCGTGA 738

QY 466 AGTCTAGGAGCAGGAAGATGGTTGGGACGTGACCGGGCCGAGGCTATGCTCCACCG 525
DB 739 AGTCTAGGAGCAGGAAGATGGTTGGGACGTGACCGGGCCGAGGCTATGCTCCACCG 798

QY 526 CCAAGTGCTGAACATCTGGGCGCTGATCTGGGACATCTCATGACATTTGGATTATCC 585
DB 799 CCAAGTGCTGAACATCTGGGCGCTGATCTGGGACATCTCATGACATTTGGATTATCC 858

QY 586 TGTACTCGTATTCGGCTCTGTGACAGTCTACATATTATGTTACAGATAATACAGAAA 645
DB 859 TGTACTCGTATTCGGCTCTGTGACAGTCTACATATTATGTTACAGATAATACAGAAA 918

QY 646 AACGGGTTACTAGTAGCGCGCCATAGCCTTGGACCTTTGCACTCCACTGTGCAATGCTG 705
DB 919 AACGGGTTACTAGTAGCGCGCCATAGCCTTGGACCTTTGCACTCCACTGTGCAATGCTG 978

QY 706 GCCCTGACGCTGGGCTGTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 765
DB 979 GCCCTGACGCTGGGCTGTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1038

QY 766 TATACCCACACCTGTCTACACTGACATTCATTAAGTG-AGTGTCTTGTGAAAAA 824
DB 1039 TATACCCACACCTGTCTACACTGACATTCATTAAGTGCACTGTCTTGTGATGCT 1098

QY 825 ACAAATAAAACCCAGAGGGGGGGCGCG 851
DB 1099 GTGACTTATCTGAGTGGGGGGCGCTG 1125

RESULT 5
US-10-240-965-263
Sequence 263, Application US/10240965
Publication No. US20030165924A1
GENERAL INFORMATION:
APPLICANT: INCYTE GENOMICS, INC.
APPLICANT: SHIFFMAN, Dov
APPLICANT: SOMOGYI, Roland
APPLICANT: LAM, Richard M.
APPLICANT: SEILHAMER, Jeffrey J.
APPLICANT: PORTER, Gordon J.
APPLICANT: MIKITA, Thomas
APPLICANT: TAI, Julie
TITLE OF INVENTION: GENES EXPRESSED IN FOAM CELL DIFFERENTIATION
FILE REFERENCE: PA-0025 PCT
CURRENT APPLICATION NUMBER: US/10/240,965
CURRENT FILING DATE: 2002-10-04
PRIOR APPLICATION NUMBER: 60/195,106
PRIOR FILING DATE: 2000-04-05
NUMBER OF SEQ ID NOS: 276
SOFTWARE: PERL Program
SEQ ID NO 263
LENGTH: 1095
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20030165924A1 988231.7
US-10-240-965-263

Query Match 63.1%; Score 699.8; DB 14; Length 1095;
Best Local Similarity 99.6%; Pred. No. 1.6e-209;
Matches 712; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 106 AGCAGGAAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAA 165
DB 381 AGCAGGAAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAA 440

QY 166 AGGAGGTCTACTGAGCAGCGTCCAGCATCCGACACACACAGCGGCGCTTGGCTCCACG 225
DB 441 AGGAGGTCTACTGAGCAGCGTCCAGCATCCGACACACACAGCGGCGCTTGGCTCCACG 500

QY 226 CAGAAACACACACTTCTCAAACTTCACTCAACACTTCTTCCCAAGCCAGCAATCCCTTCCAAAGGT 285
DB 501 CAGAAACACACACTTCTCAAACTTCACTCAACACTTCTTCCCAAGCCAGCAATCCCTTCCAAAGGT 560

QY 286 ACAAGGAGAAATAGAGGTGGTGTCTGGGGGACACCCCGGACACCATCTTCCAAAGGT 345
DB 561 ACAAGGAGAAATAGAGGTGGTGTCTGGGGGACACCCCGGACACCATCTTCCAAAGGT 620

QY 346 CCACCGTGATCAACATCCACAGCAGACCTCCGTGCGCCGACCATCTGCTGGTCCCTGT 405
DB 621 CCACCGTGATCAACATCCACAGCAGACCTCCGTGCGCCGACCATCTGCTGGTCCCTGT 680

QY 406 TCAACACCTCTTCTTGAACCTGGTGTCTGGGCTTCATAGCATTCGCTACTCCGTA 465
DB 681 TCAACACCTCTTCTTGAACCTGGTGTCTGGGCTTCATAGCATTCGCTACTCCGTA 740

QY 466 AGTCTAGGAGCAGGAAGATGGTTGGGACGTGACCGGGCCGAGGCTATGCTCCACCG 525
DB 741 AGTCTAGGAGCAGGAAGATGGTTGGGACGTGACCGGGCCGAGGCTATGCTCCACCG 800

QY 526 CCAAGTGCTGAACATCTGGGCGCTGATCTGGGACATTCCTCATGACCATTTGGATTATCC 585
DB 801 CCAAGTGCTGAACATCTGGGCGCTGATCTGGGACATTCCTCATGACCATTTGGATTATCC 860

QY 586 TGTACTGGTATTCGGCTCTGTGACAGTCTACCATATTATGTTACAGATAATACAGAAA 645

Db 861 TGTACTGGTATTCGGCTCTGTGACAGCTCTACCATATTTATGTTACAGATAATACAGGAAA 920
 QY 646 AACGGGGTTACTAGTAGCCGCCCATAGCCTGCAACCTTTGCACTCCCACTGTGCAATGCTG 705
 Db 921 AACGGGGTTACTAGTAGCCGCCCATAGCCTGCAACCTTTGCACTCCCACTGTGCAATGCTG 980
 QY 706 GGCCTGCAGCTGGGGCTGTGGCCCTGGCCCTTGGTCTGCTGCCCTAGATACAGCAGTT 765
 Db 981 GGCCTGCAGCTGGGGCTGTGGCCCTGGCCCTTGGTCTGCTGCCCTAGATACAGCAGTT 1040
 QY 766 TATACCCACACACTGCTACACTGACATTCATAAAGTG-ACGTGCTTGTGAAA 819
 Db 1041 TATACCCACACACTGCTACAGTGTCTATCAATAAAGTGACGTGCTTGTGAAA 1095

RESULT 6
 US-09-920-300A-1743
 ; Sequence 1743, Application US/09920300A
 ; Patent No. US20020136728A1
 ; GENERAL INFORMATION:
 ; APPLICANT: King, Gordon E.
 ; APPLICANT: Meagher, Madeleine Joy
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Secrist, Heather
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
 ; FILE REFERENCE: 210121.547
 ; CURRENT APPLICATION NUMBER: US/09/920.300A
 ; CURRENT FILING DATE: 2001-07-31
 ; NUMBER OF SEQ ID NOS: 1789
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1743
 ; LENGTH: 842
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-920-300A-1743

Query Match 62.8%; Score 696.2; DB 9; Length 842;
 Best Local Similarity 99.4%; Pred. No. 1.9e-209;
 Matches 709; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
 QY 106 AGCAGGAATAGAACTTAAAGAAATACACACTTCTGAGAACTTGAACACAGCGGAA 165
 Db 130 AGCAGGAATAGAACTTAAAGAAATACACACTTCTGAGAACTTGAACACAGCGGAA 189
 QY 166 AGGAGGTCTCACTGAGCACCCTCCAGCATCCGGACACACAGCGGCCCTTCGCTCCAG 225
 Db 190 AGGAGGTCTCACTGAGCACCCTCCAGCATCCGGACACACAGCGGCCCTTCGCTCCAG 249
 QY 226 CAGAAACACACTTCTCAACCTTCACTCAACACTTCTTCCCAAGCCAGAGATGC 285
 Db 250 CAGAAACACACTTCTCAACCTTCACTCAACACTTCTTCCCAAGCCAGAGATGC 309
 QY 286 ACAAGGAGGAACATGAGTGGCTGTGTGGGGGACACCCCGACACCATCTTCCAAGGT 345
 Db 310 ACAAGGAGGAACATGAGTGGCTGTGTGGGGGACACCCCGACACCATCTTCCAAGGT 369
 QY 346 CCACCGTGATCAACATCCACAGCAGACCTCCGTGCCGACCATGTGCTGGTCCCTGT 405
 Db 370 CCACCGTGATCAACATCCACAGCAGACCTCCGTGCCGACCATGTGCTGGTCCCTGT 429
 QY 406 TCAACACCCCTTCTTGAACCTGTGTGTGGGCTTCATAGCATTCGCGCTACTCCGTGA 465
 Db 430 TCAACACCCCTTCTTGAACCTGTGTGTGGGCTTCATAGCATTCGCGCTACTCCGTGA 489
 QY 466 AGTCTAGGACAGGAAGATGTTGGCGAGCTGACCGGGGCCAGGCCCTATGCTCCACCG 525
 Db 490 AGTCTAGGACAGGAAGATGTTGGCGAGCTGACCGGGGCCAGGCCCTATGCTCCACCG 549
 QY 526 CCAAGTGCCTGAACATCTGGGCCCTGATTTCTGGGCATCTCTATGACCATTTGATTCATCC 585
 Db 550 CCAAGTGCCTGAACATCTGGGCCCTGATTTCTGGGCATCTCTATGACCATTTGATTCATCC 609

QY 586 TGTACTGGTATTCGGCTCTGTGACAGCTCTACCATATTTATGTTACAGATAATACAGGAAA 645
 Db 610 TGTACTGGTATTCGGCTCTGTGACAGCTCTACCATATTTATGTTACAGATAATACAGGAAA 669
 QY 646 AACGGGGTTACTAGTAGCCGCCCATAGCCTGCAACCTTTGCACTCCCACTGTGCAATGCTG 705
 Db 670 AACGGGGTTACTAGTAGCCGCCCATAGCCTGCAACCTTTGCACTCCCACTGTGCAATGCTG 729
 QY 706 GGCCTGCAGCTGGGGCTGTGGCCCTGGCCCTTGGTCTGCTGCCCTAGATACAGCAGTT 765
 Db 730 GGCCTGCAGCTGGGGCTGTGGCCCTGGCCCTTGGTCTGCTGCCCTAGATACAGCAGTT 789
 QY 766 TATACCCACACACTGCTACACTGACATTCATAAAGTG-ACGTGCTTGTGAAA 817
 Db 790 TATACCCACACACTGCTACAGTGTCTATCAATAAAGTGACGTGCTTGTGAAA 842

RESULT 7
 US-10-033-528-1743
 ; Sequence 1743, Application US/10033528
 ; Publication No. US20020131971A1
 ; GENERAL INFORMATION:
 ; APPLICANT: King, Gordon E.
 ; APPLICANT: Meagher, Madeleine Joy
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Secrist, Heather
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
 ; FILE REFERENCE: 210121.547C1
 ; CURRENT APPLICATION NUMBER: US/10/033.528
 ; CURRENT FILING DATE: 2001-12-26
 ; NUMBER OF SEQ ID NOS: 1896
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1743
 ; LENGTH: 842
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-033-528-1743

Query Match 62.8%; Score 696.2; DB 13; Length 842;
 Best Local Similarity 99.4%; Pred. No. 1.9e-209;
 Matches 709; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
 QY 106 AGCAGGAATAGAACTTAAAGAAATACACACTTCTGAGAACTTGAACACAGCGGAA 165
 Db 130 AGCAGGAATAGAACTTAAAGAAATACACACTTCTGAGAACTTGAACACAGCGGAA 189
 QY 166 AGGAGGTCTCACTGAGCACCCTCCAGCATCCGGACACACAGCGGCCCTTCGCTCCAG 225
 Db 190 AGGAGGTCTCACTGAGCACCCTCCAGCATCCGGACACACAGCGGCCCTTCGCTCCAG 249
 QY 226 CAGAAACACACTTCTCAACCTTCACTCAACACTTCTTCCCAAGCCAGAGATGC 285
 Db 250 CAGAAACACACTTCTCAACCTTCACTCAACACTTCTTCCCAAGCCAGAGATGC 309
 QY 286 ACAAGGAGGAACATGAGTGGCTGTGTGGGGGACACCCCGACACCATCTTCCAAGGT 345
 Db 310 ACAAGGAGGAACATGAGTGGCTGTGTGGGGGACACCCCGACACCATCTTCCAAGGT 369
 QY 346 CCACCGTGATCAACATCCACAGCAGACCTCCGTGCCGACCATGTGCTGGTCCCTGT 405
 Db 370 CCACCGTGATCAACATCCACAGCAGACCTCCGTGCCGACCATGTGCTGGTCCCTGT 429
 QY 406 TCAACACCCCTTCTTGAACCTGTGTGTGGGCTTCATAGCATTCGCGCTACTCCGTGA 465
 Db 430 TCAACACCCCTTCTTGAACCTGTGTGTGGGCTTCATAGCATTCGCGCTACTCCGTGA 489
 QY 466 AGTCTAGGACAGGAAGATGTTGGCGAGCTGACCGGGGCCAGGCCCTATGCTCCACCG 525
 Db 490 AGTCTAGGACAGGAAGATGTTGGCGAGCTGACCGGGGCCAGGCCCTATGCTCCACCG 549
 QY 526 CCAAGTGCCTGAACATCTGGGCCCTGATTTCTGGGCATCTCTATGACCATTTGATTCATCC 585

Db 550 CCAAGTGCCTGAAACATCTGGGCCCTGATTTCTGGGCATCCTCATGACCAATTGGATTTCATCC 609
 QY 586 TGTACTGGTATTGGCTCTGTGACAGTCTACCATATTATTTACAGATAATACAGGAA 645
 Db 610 TGTCACTGGTATTGGCTCTGTGACAGTCTACCATATTATTTACAGATAATACAGGAA 669
 QY 646 AACGGGTTACTAGTAGCGGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTG 705
 Db 670 AACGGGTTACTAGTAGCGGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTG 729
 QY 706 GCCCTGACAGTGGGCTGTTGCCCTGCCCCCTGGTCTGCTCCCTAGATACAGAGTT 765
 Db 730 GCCCTGACAGTGGGCTGTTGCCCTGCCCCCTGGTCTGCTCCCTAGATACAGAGTT 789
 QY 766 TATACCCACACACCTGTCTACACTGACATTCATTAAGTG-ACGTGCTTGTGA 817
 Db 790 TATACCCACACACCTGTCTACAGTGTCAATTAAGTGACAGTGTGTTGTA 842

RESULT 8

US-10-099-926-1743

; Sequence 1743, Application US/10099926

; Publication No. US200301660641

; GENERAL INFORMATION:

; APPLICANT: King, Gordon E.

; APPLICANT: Mesgher, Madeleine Joy

; APPLICANT: Xu, Jiangchun

; APPLICANT: Secrist, Heather

; APPLICANT: Jiang, Yugu

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.547C2

; CURRENT APPLICATION NUMBER: US/10/099,926

; CURRENT FILING DATE: 2002-03-17

; NUMBER OF SEQ ID NOS: 1982

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1743

; LENGTH: 842

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-099-926-1743

Query Match 62.8%; Score 696.2; DB 14; Length 842;
 Best Local Similarity 99.4%; Pred. No. 1.9e-208;
 Matches 709; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 106 AGCAGGAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAACGACAGGGGAA 165
 Db 130 AGCAGGAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAACGACAGGGGAA 189
 QY 166 AGGAGTCTCACTGAGCAGCGTCCAGCATCCGACACACAGCGGCCCTTGGCTCCACG 225
 Db 190 AGGAGTCTCACTGAGCAGCGTCCAGCATCCGACACACAGCGGCCCTTGGCTCCACG 249
 QY 226 CAGAAACACACTTCTCAACCTTCACTCAACACTTCTTCCCAAGCCAGAGATGC 285
 Db 250 CAGAAACACACTTCTCAACCTTCACTCAACACTTCTTCCCAAGCCAGAGATGC 309
 QY 286 ACAAGGAGAAATAGAGTGGCTGTGTGGGGGCACCCCGCAGCACCATCTTCCAAAGT 345
 Db 310 ACAAGGAGAAATAGAGTGGCTGTGTGGGGGCACCCCGCAGCACCATCTTCCAAAGT 369
 QY 346 CCACCGTGATCAACATCCAGGAGACCTCGTGCCCGACCATGTGCTGGTCCCTGT 405
 Db 370 CCACCGTGATCAACATCCAGGAGACCTCGTGCCCGACCATGTGCTGGTCCCTGT 429
 QY 406 TCAACACCTCTTCTTGAATCTGTGTCTGGGCTTCATAGCATTCGCTTACTCCGTTGA 465
 Db 430 TCAACACCTCTTCTTGAATCTGTGTCTGGGCTTCATAGCATTCGCTTACTCCGTTGA 489
 QY 466 AGTCTAGGACAGGAATGGTGGGAGCGTACCGGGGCCAGGCTATGCTTCCACCG 525

Db 490 AGTCTAGGACAGGAATGGTGGCGAGTGACCGGGGCCAGGGCTATGCTTCCACCG 549
 QY 526 CCAAGTGCCTGAAACATCTGGGCCCTGATTTCTGGGCATCCTCATGACCAATTGGATTTCATCC 585
 Db 550 CCAAGTGCCTGAAACATCTGGGCCCTGATTTCTGGGCATCCTCATGACCAATTGGATTTCATCC 609
 QY 586 TGTACTGGTATTGGCTCTGTGACAGTCTACCATATTATTTACAGATAATACAGGAA 645
 Db 610 TGTCACTGGTATTGGCTCTGTGACAGTCTACCATATTATTTACAGATAATACAGGAA 669
 QY 646 AACGGGTTACTAGTAGCGGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTG 705
 Db 670 AACGGGTTACTAGTAGCGGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTG 729
 QY 706 GCCCTGACAGTGGGCTGTTGCCCTGCCCCCTGGTCTGCTCCCTAGATACAGAGTT 765
 Db 730 GCCCTGACAGTGGGCTGTTGCCCTGCCCCCTGGTCTGCTCCCTAGATACAGAGTT 789
 QY 766 TATACCCACACACCTGTCTACACTGACATTCATTAAGTG-ACGTGCTTGTGA 817
 Db 790 TATACCCACACACCTGTCTACAGTGTCAATTAAGTGACAGTGTGTTGTA 842

RESULT 9

US-09-880-107-2108

; Sequence 2108, Application US/09880107

; Patent No. US20020142981A1

; GENERAL INFORMATION:

; APPLICANT: Horne, Darcie T.

; APPLICANT: Vockley, Joseph G.

; APPLICANT: Scherf, Uwe

; APPLICANT: Gene Logic, Inc.

; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer

; FILE REFERENCE: 44921-5028-WO

; CURRENT APPLICATION NUMBER: US/09/880,107

; CURRENT FILING DATE: 2001-06-14

; PRIOR APPLICATION NUMBER: US 60/211,379

; PRIOR FILING DATE: 2000-06-14

; PRIOR APPLICATION NUMBER: US 60/237,054

; PRIOR FILING DATE: 2000-10-02

; NUMBER OF SEQ ID NOS: 3950

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2108

; LENGTH: 853

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: Genbank Accession No. US20020142981A1 J04164

US-09-880-107-2108

Query Match 62.8%; Score 696.2; DB 9; Length 853;
 Best Local Similarity 99.4%; Pred. No. 1.9e-208;
 Matches 709; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 106 AGCAGGAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAACGACAGGGGAA 165
 Db 141 AGCAGGAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAACGACAGGGGAA 200
 QY 166 AGGAGTCTCACTGAGCAGCGTCCAGCATCCGACACACAGCGGCCCTTGGCTCCACG 225
 Db 201 AGGAGTCTCACTGAGCAGCGTCCAGCATCCGACACACAGCGGCCCTTGGCTCCACG 260
 QY 226 CAGAAACACACTTCTCAACCTTCACTCAACACTTCTTCCCAAGCCAGAGATGC 285
 Db 261 CAGAAACACACTTCTCAACCTTCACTCAACACTTCTTCCCAAGCCAGAGATGC 320
 QY 286 ACAAGGAGAAATAGAGTGGCTGTGCTGGGGGCACCCCGCAGCACCATCTTCCAAAGT 345
 Db 321 ACAAGGAGAAATAGAGTGGCTGTGCTGGGGGCACCCCGCAGCACCATCTTCCAAAGT 380
 QY 346 CCACCGTGATCAACATCCAGGAGACCTCGTGCCCGACCATGTGCTGGTCCCTGT 405
 Db 381 CCACCGTGATCAACATCCAGGAGACCTCGTGCCCGACCATGTGCTGGTCCCTGT 440

QY 406 TCAACACCCCTCTTCTGAACTGGTGTCTGTGGCTTCATAGCAATCGCCTACTCCGTGA 465
 DB 441 TCAACACCCCTCTTCTGAACTGGTGTCTGTGGCTTCATAGCAATCGCCTACTCCGTGA 500
 QY 466 AGTCTAGGACAGAGAGATGGTGGCGACGTGACCGGGGCCCGAGCCCTATGCTCCACCG 525
 DB 501 AGTCTAGGACAGAGAGATGGTGGCGACGTGACCGGGGCCCGAGCCCTATGCTCCACCG 560
 QY 526 CCAAGTGCCTGAACATCTGGCCCTGATTTCTGGGATCTCTATGACCAATGGATTATCC 585
 DB 561 CCAAGTGCCTGAACATCTGGCCCTGATTTCTGGGATCTCTATGACCAATGGATTATCC 620
 QY 586 TGTACTGTGATTCGGCTCTGTGACAGTCTACCATATTTATGTTACAGATAATACAGGAAA 645
 DB 621 TGTACTGTGATTCGGCTCTGTGACAGTCTACCATATTTATGTTACAGATAATACAGGAAA 680
 QY 646 AACGGGGTTACTAGTACCGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTG 705
 DB 681 AACGGGGTTACTAGTACCGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTG 740
 QY 706 GCCTTCACGCTGGGCTGTGGCCCTGCGCCCTTGGTCTCGCCCTAGATACAGCAGTT 765
 DB 741 GCCTTCACGCTGGGCTGTGGCCCTGCGCCCTTGGTCTCGCCCTAGATACAGCAGTT 800
 QY 766 TATACCCACACACTGTCTACACTGACATTCATTAAGTG-ACGTGCTTGTGA 817
 DB 801 TATACCCACACACTGTCTACACTGACATTCATTAAGTG-ACGTGCTTGTGA 853

RESULT 10
 US-09-960-706-626
 ; Sequence 626, Application US/09960706
 ; Publication No. US20030134280A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mungler, William E.
 ; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplasia
 ; FILE REFERENCE: 44921-5029-01US
 ; CURRENT APPLICATION NUMBER: US/09/960,706
 ; CURRENT FILING DATE: 2001-09-24
 ; PRIOR APPLICATION NUMBER: 60/223,323
 ; PRIOR FILING DATE: 2000-08-07
 ; PRIOR APPLICATION NUMBER: 09/873,319
 ; PRIOR FILING DATE: 2001-06-05
 ; NUMBER OF SEQ ID NOS: 1124
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 626
 ; LENGTH: 853
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: Genbank Accession No. US20030134280A1 J04164
 US-09-960-706-626

Query Match 62.8%; Score 696.2; DB 10; Length 853;
 Best Local Similarity 99.4%; Pred. No. 1.9e-208;
 Matches 709; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
 QY 106 AGCAGGAAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAA 165
 DB 141 AGCAGGAAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAA 200
 QY 166 AGAGGTCTCACTGAGCACCGTCCAGATCCGGACACACAGCGGCCCTTGGCTCCACG 225
 DB 201 AGAGGTCTCACTGAGCACCGTCCAGATCCGGACACACAGCGGCCCTTGGCTCCACG 260
 QY 226 CAGAAACACACACTTCTCAACCTTCACTCAACACTTCTTCCCAAGCCAGAGATGC 285
 DB 261 CAGAAACACACACTTCTCAACCTTCACTCAACACTTCTTCCCAAGCCAGAGATGC 320
 QY 286 ACAAGGAGAACATGAGGTGGCTGTGGGGGACACCCCGACAGACCATCTTCCAGGT 345

DB 321 ACAAGGAGAACATGAGGTGGCTGTGGGGCACCCCCAGCAGCACTCTTCCAAGT 380
 QY 346 CCAACCGTGTCAACATCCACAGAGACCTCCGTGCGCCGACCACTGTCTGTGGTCCCTGT 405
 DB 381 CCAACCGTGTCAACATCCACAGAGACCTCCGTGCGCCGACCACTGTCTGTGGTCCCTGT 440
 QY 406 TCAACACCCCTCTTCTTGAACCTGGTGTCTGTGGGCTTCATAGCAATTCGCTACTCCGTGA 465
 DB 441 TCAACACCCCTCTTCTTGAACCTGGTGTCTGTGGGCTTCATAGCAATTCGCTACTCCGTGA 500
 QY 466 AGTCTAGGACAGAGAGATGGTGGCGACGTGACCGGGGCCCGAGCCCTATGCTCCACCG 525
 DB 501 AGTCTAGGACAGAGAGATGGTGGCGACGTGACCGGGGCCCGAGCCCTATGCTCCACCG 560
 QY 526 CCAAGTGCCTGAACATCTGGCCCTGATTTCTGGGATCTCTATGACCAATGGATTATCC 585
 DB 561 CCAAGTGCCTGAACATCTGGCCCTGATTTCTGGGATCTCTATGACCAATGGATTATCC 620
 QY 586 TGTACTGTGATTCGGCTCTGTGACAGTCTACCATATTTATGTTACAGATAATACAGGAAA 645
 DB 621 TGTACTGTGATTCGGCTCTGTGACAGTCTACCATATTTATGTTACAGATAATACAGGAAA 680
 QY 646 AACGGGGTTACTAGTACCGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTG 705
 DB 681 AACGGGGTTACTAGTACCGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTG 740
 QY 706 GCCTTCACGCTGGGCTGTGGCCCTGCGCCCTTGGTCTCGCCCTAGATACAGCAGTT 765
 DB 741 GCCTTCACGCTGGGCTGTGGCCCTGCGCCCTTGGTCTCGCCCTAGATACAGCAGTT 800
 QY 766 TATACCCACACACTGTCTACACTGACATTCATTAAGTG-ACGTGCTTGTGA 817
 DB 801 TATACCCACACACTGTCTACACTGACATTCATTAAGTG-ACGTGCTTGTGA 853

RESULT 11
 US-09-873-319-392
 ; Sequence 392, Application US/09873319A
 ; Publication No. US20030134324A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mungler, William E.
 ; APPLICANT: Kulkarni, Prakash
 ; APPLICANT: Getzenberg, Robert H.
 ; APPLICANT: Waga, Iwao
 ; APPLICANT: Yamamoto, Jun
 ; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic
 ; FILE REFERENCE: 44921-5029-US
 ; CURRENT APPLICATION NUMBER: US/09/873,319A
 ; CURRENT FILING DATE: 2001-06-05
 ; EARLIER APPLICATION NUMBER: US 60/223,323
 ; EARLIER FILING DATE: 2000-08-07
 ; NUMBER OF SEQ ID NOS: 755
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 392
 ; LENGTH: 853
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: Genbank Accession No. US20030134324A1 J04164
 US-09-873-319-392

Query Match 62.8%; Score 696.2; DB 10; Length 853;
 Best Local Similarity 99.4%; Pred. No. 1.9e-208;
 Matches 709; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
 QY 106 AGCAGGAAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAA 165
 DB 141 AGCAGGAAATAGAACTTAAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAA 200
 QY 166 AGAGGTCTCACTGAGCACCGTCCAGATCCGGACACACAGCGGCCCTTCCCTCCACG 225
 DB 201 AGAGGTCTCACTGAGCACCGTCCAGATCCGGACACACAGCGGCCCTTCCCTCCACG 260

Job time : 261.529 secs

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MM01-0292RNM
; CURRENT APPLICATION NUMBER: US/10/301,822
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 80
; LENGTH: 647
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (111)...(488)
; US-10-301-822-80

Query Match 56.3%; Score 624.4; DB 14; Length 647;
 Best Local Similarity 98.9%; Pred. No. 7e-186;
 Matches 639; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

Qy	173	CTCACTGAGCAGCGTCCGAGCATCCGAGACACACAGCGGCGCTTCGCTCCACGCAGAAA	232
Db	2	CTCACTGAGCAGCGTCCGAGCATCCGAGACACACAGCGGCGCTTCGCTCCACGCAGAAA	61
Qy	233	CCACACTTCTCAAACTTCACACTCAACACTTCCTTCCCAAAGCCAGAGATGCACAAAGGA	292
Db	62	CCACACTTCTCACTTCTCACTCAACACTTCCTTCCCAAAGCCAGAGATGCACAAAGGA	121
Qy	293	GGAAATGAGGTGGTGTGCTGGGGGACACCCCGAGCACCATTCTTCRAAGTCCACCGT	352
Db	122	GGAAATGAGGTGGTGTGCTGGGGGACACCCCGAGCACCATTCTTCRAAGTCCACCGT	181
Qy	353	GATCAACATCCACAGGAGACCTCCGTCGCCGACCATGTGCTGGTCCCTGTTCAACAC	412
Db	182	GATTAAATCCACAGGAGACCTCCGTCGCCGACCATGTGCTGGTCCCTGTTCAACAC	241
Qy	413	CCTCTTCTTGAAGTGTGCTGGCTTCATAGCATTCGCTACTCCGTTGAGTCTAG	472
Db	242	CCTCTTCTTGAAGTGTGCTGGCTTCATAGCATTCGCTACTCCGTTGAGTCTAG	301
Qy	473	GGACAGGAAGATGGTTGGGACGTGACCGGGGCCAGGCGCTATGCTCCACCGCAAAGTG	532
Db	302	GGACAGGAAGATGGTTGGGACGTGACCGGGGCCAGGCGCTATGCTCCACCGCAAAGTG	361
Qy	533	CCTGAACATCTGGGCGCTGATTCCTGGGATCCTCATGACCATTTGGATTATCCTGTTACT	592
Db	362	CCTGAACATCTGGGCGCTGATTCCTGGGATCCTCATGACCATTTGGATTATCCTGTTACT	421
Qy	593	GGTATTCGGCTCTGTACAGTCTACCATATTTATGTTACAGATAACAGGAAAAACGGGG	652
Db	422	GGTATTCGGCTCTGTACAGTCTACCATATTTATGTTACAGATAACAGGAAAAACGGGG	481
Qy	653	TTACTAGTAGCCGCCCATAGCCTGCAACCTTTGCACTTCCACTGTGCAATGTGCGCCCTGC	712
Db	482	TTACTAGTAGCCGCCCATAGCCTGCAACCTTTGCACTTCCACTGTGCAATGTGCGCCCTGC	541
Qy	713	ACGCTGGGGCTGTGCGCCCTGCGCCCTTGGTCCCTGAGATACAGCAGTTTATACCC	772
Db	542	ACGCTGGGGCTGTGCGCCCTGCGCCCTTGGTCCCTGAGATACAGCAGTTTATACCC	601
Qy	773	ACACACCTGTCTACACTGACATTCATTAAGTG-ACGTGTTGTGA	817
Db	602	ACACACCTGTCTACAGTGTATTCAATAAAGTGACGCTGCTTGTGA	647

Search completed: March 12, 2004; 06:15:10

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 10, 2004, 09:18:38 ; Search time 1836.07 Seconds

(without alignments)
18036.975 Million cell updates/sec

Title: US-10-084-817-37

Perfect score: 1109

Sequence: 1 cctgaccaggagacactgg.....acanaattccacctatccg 1109

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estnu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_hic:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_hic:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: em_gss_hum:*
- 18: em_gss_inv:*
- 19: em_gss_pln:*
- 20: em_gss_vit:*
- 21: em_gss_fun:*
- 22: em_gss_mam:*
- 23: em_gss_mus:*
- 24: em_gss_pro:*
- 25: em_gss_fod:*
- 26: em_gss_phg:*
- 27: em_gss_vrl:*
- 28: gb_gss1:*
- 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	726.8	65.5	841	12	BI909898 603070530
2	689.8	62.2	727	14	CD174424 AGENCOURT
3	685	61.8	835	13	BU603243 AGENCOURT
4	682.8	61.6	728	14	CD174239 AGENCOURT

5	682.2	61.5	718	14	CF780624
6	678.2	61.2	717	14	CD175008
7	678	61.1	726	14	CD175837
8	676.2	61.0	718	14	CD172759
9	672.8	60.7	726	14	CD176014
10	671.8	60.6	721	14	CD175245
11	668.8	60.3	727	9	AV759074
12	668	60.2	704	14	CD173064
13	667.2	60.2	715	14	CD173537
14	666.6	60.1	713	14	CD173073
15	666.4	60.1	712	14	CD175286
16	665.4	60.0	710	14	CD172976
17	665	60.0	710	14	CD176886
18	665	60.0	710	14	CD247128
19	664.6	59.9	691	13	BQ219055
20	663	59.8	699	14	CD250223
21	662	59.7	699	12	BG719484
22	661.2	59.6	709	14	CD175451
23	661	59.6	715	14	CD172777
24	661	59.6	726	14	CD249234
25	660.6	59.6	713	14	CD174839
26	660	59.5	711	14	CD247389
27	659	59.4	738	13	BUS37783
28	658.2	59.4	707	14	CD173531
29	657.6	59.3	677	14	CA421204
30	655.8	59.1	691	13	BUS58290
31	653.4	58.9	700	14	CD247008
32	653	58.9	707	14	CD173309
33	652.6	58.8	793	9	AV758898
34	651.8	58.8	703	14	CD175270
35	651.6	58.8	670	12	BM995109
36	651	58.7	699	14	CD385219
37	650.2	58.6	669	12	BG534713
38	649.6	58.6	697	14	CD702719
39	649.4	58.6	704	14	CD248465
40	649	58.5	712	14	CD512187
41	647	58.3	895	13	BQ228236
42	646.8	58.3	762	10	BF665518
43	641.2	57.8	662	9	AV759976
44	639.6	57.7	706	14	CA429526
45	639.2	57.6	673	14	CB151066

ALIGNMENTS

RESULT 1
BI909898
LOCUS
DEFINITION 603070530F1 NIH_MGC_118 Homo sapiens CDNA clone IMAGE:5219479 5',
mRNA sequence.
ACCESSION BI909898
VERSION BI909898.1 GI:16173196
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 841)
NIH-MGC http://mgc.ncbi.nlm.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: csgbbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LHAM11551 row: 1 column: 08
High quality sequence stop: 824.

BI909898 841 bp mRNA linear EST 16-OCT-2001
603070530F1 NIH_MGC_118 Homo sapiens CDNA clone IMAGE:5219479 5',
mRNA sequence.

FEATURES source Location/Qualifiers

1. 841 /organism="Homo sapiens" /mol_type="rRNA" /db_xref="taxon:9606" /clone="IMAGE:5219479" /tissue_type="leukocyte" /lab_host="DH103" /clone_lib="NIH_MGC 118" /note="Vector: pCMV-SPT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 65.5%; Score 726.8; DB 12; Length 841; Best Local Similarity 99.3%; Pred. No. 5.6e-137; Matches 761; Conservative 0; Mismatches 2; Indels 3; Gaps 3;

QY 1 COTGACAGGAGACTGGAGGTTAGTCCCAACCCGACAGAGGAGGACTGCAGC 60

Db 59 COTGACAGGAGACTGGAGGTTAGTCCCAACCCGACAGAGGAGGACTGCAGC 118

QY 61 CTGAGGAAGACAGAGATTTCAGGAGAGAGGCTCGGACAAAGTGACAGGAATAGAAA 120

Db 119 CTGAGGAAGACAGAGATTTCAGGAGAGAGGCTCGGACAAAGTGACAGGAATAGAAA 178

QY 121 CTTAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAAAGGAGTCTCACTGA 180

Db 179 CTTAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAAAGGAGTCTCACTGA 238

QY 181 GCACCGTCCAGCATCCGACACACAGCGGGCTTCGCTCCAGCGAAGAAACACACTT 240

Db 239 GCACCGTCCAGCATCCGACACACAGCGGGCTTCGCTCCAGCGAAGAAACACACTT 298

QY 241 CTCAAACCTTCACTCAACACTTCTCCCAAGCAGAGAGATGCAACAGGAGAAACATG 300

Db 299 CTCAAACCTTCACTCAACACTTCTCCCAAGCAGAGAGATGCAACAGGAGAAACATG 358

QY 301 AGGTGGCTGTGTGGGGACCCCGACAGACCATCTTCAAGTCCACCGTGATCAACA 360

Db 359 AGGTGGCTGTGTGGGGACCCCGACAGACCATCTTCAAGTCCACCGTGATCAACA 418

QY 361 TCCACAGCAGACCTCGTGCCGACCATGTCGTCTGCTCCCTGTTCAACACCTCTTCT 420

Db 419 TCCACAGCAGACCTCGTGCCGACCATGTCGTCTGCTCCCTGTTCAACACCTCTTCT 478

QY 421 TGAACCTGGTGTCTGTGGGTTTCATAGCATTCGCTTACTCGTGAAGTCTAGGGACAGA 480

Db 479 TGAACCTGGTGTCTGTGGGTTTCATAGCATTCGCTTACTCGTGAAGTCTAGGGACAGA 538

QY 481 AGATGTTGGCGACGTACCGGGCCCGACGCTATGCTCCACCGCAAGTCCCTGGAACA 540

Db 539 AGATGTTGGCGACGTACCGGGCCCGACGCTATGCTCCACCGCAAGTCCCTGGAACA 598

QY 541 TCTGGGCCCTGATTCCTGGGCATCCTCATGACCATTTGATTCATCCTGTACTGGTATTTCG 600

Db 599 TCTGGGCCCTGATTCCTGGGCATCCTCATGACCATTTGATTCATCCTGTACTGGTATTTCG 658

QY 601 GCTCTGTGACAGTACCATATTATGTTACAGATAATACAGGAAACGGGGTACTAGT 660

Db 659 GCTCTGTGACAGTACCATATTATGTTACAGATAATACAGGAAACGGGGTACTAGT 718

QY 661 AGCGGCCCATAGCTGCAACTTTCGACTCCACTG-TGCAATGCTGGCCCTGCACGCTGG 719

Db 719 AGCGG-CCATAGCTGCAACTTTCGACTCCACTGTTGCAATGCTGGCCCTGCACGCTGG 777

QY 720 GGCTGTTGGCCCTTGGCTTGGCTTGGCCCTAGATACAGCAGTT 765

Db 778 GGCTGTTGGCCCTG-CCCTTGGTCTTGCCTTAGATACAGCAGTT 822

RESULT 2

CD174424 LOCUS 727 bp mRNA linear EST 19-MAY-2003

DEFINITION AGENCOURT_13973296 NIH_MGC_172 Homo sapiens cDNA 5', mRNA sequence.

ACCESSION CD174424

VERSION CD174424.1 GI:30857965

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 727)

NIH-MGC http://mgi.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Dr. Jamie Thompson, University of WI

cDNA Library Preparation: Gina Zastrow-Hayes

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov

Plate: NDRM48 row: p column: 16

High quality sequence start: 9

High quality sequence stop: 727.

Location/Qualifiers

1. 727 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /tissue_type="embryonic stem cells, WA01, passage 38" /lab_host="NIH_MGC 172" /clone_lib="NIH_MGC 172" /note="Vector: pDONR201; Site 1: attP2; Site 2: attP1; LIBR. PRIMING - oligo dT; METHOD - full-length enriched; Embryonic Stem Cells H1; LIBR PROVIDER - Bradfield"

ORIGIN

Query Match 62.2%; Score 689.8; DB 14; Length 727; Best Local Similarity 98.1%; Pred. No. 1.7e-133; Indels 1; Gaps 1; Matches 708; Conservative 0; Mismatches 13;

QY 117 GAAACTTAAGAGAAATACACACTTCTGAGAAACTGAAACGACAGGGGAAAGAGGTCTCA 176

Db 5 GAAGGTAAGAAANATACACACTTCTGAGAAACTGAAACGACAGGGGAAAGAGGTCTCA 64

QY 177 CTGAGCACCGTCCAGCATCCGACACACAGCGGGCCCTTCGCTCCACGAGAAACCCAC 236

Db 65 CTGAGCACCGTCCAGCATCCGACACACAGCGGGCCCTTCGCTCCACGAGAAACCCAC 124

QY 237 ACTTCTCAAACTTCACTCAACACTTCTTCTCCCAAGCCAGAGATGCACAAGGAGAA 296

Db 125 ACTTCTCAAACTTCACTCAACACTTCTTCTCCCAAGCCAGAGATGCACAAGGAGAA 184

QY 297 CATGAGGTGGCTGTGTGGGGGACCCCGACGACCATCTTCCAAAGTCCACCGTGATC 356

Db 185 CATGAGGTGGCTGTGTGGGGGACCCCGACGACCATCTTCCAAAGTCCACCGTGATC 244

QY 357 AACATCCACAGCAGACCTCCGTGCGCGACCATGTCTGTCTGCTCCCTGTTCACACCTTC 416

Db 245 AACATCCACAGCAGACCTCCGTGCGCGACCATGTCTGTCTGCTCCCTGTTCACACCTTC 304

QY 417 TTCTTGAACTGGTGTGTCTGGGGTTTCATAGCATTCGCTACTCGCTAGGCTAGGGAC 476

Db 305 TTCTTGAACTGGTGTGTCTGGGGTTTCATAGCATTCGCTACTCGCTAGGCTAGGGAC 364

QY 477 AGGAAGATGTTGGCGACGTGACCGGGCCCGAGGCTATGCTCCACCCCAAGTGCCTG 536


```

365 AGAAGATGTTGGCAGCTACCGGGCCCGAGGCTATGCTCCACCGCCAAAGTGCCTG 424
537 AACATCTGGGCGCCGATTCGGGATCCTCATGACCATGGATTTCATCTGTTACTGGTA 596
425 AACATCTGGGCGCCGATTCGGGATCCTCATGACCATGGATTTCATCTGTTACTGGTA 484
597 TTCGGCTCTGTGACAGTCTACCATATTATGTTACAGATAAATACAGAAACGCGGTTAC 656
485 TTCGGCTCTGTGACAGTCTACCATATTATGTTACAGATAAATACAGAAACGCGGTTAC 544
657 TAGTAGCGCGCCATAGCTGCAACTTGTGCACTCCACTGTGCAATGCTGGCCCTGCACGC 716
545 TAGTAGCGCGCCATAGCTGCAACTTGTGCACTCCACTGTGCAATGCTGGCCCTGCACGC 604
717 TGGGGCTGTGGCCCTGCCCTTGGTCTGCTCCCTAGATACAGAGTTTATACCCACAC 776
605 TGGGGCTGTGGCCCTGCCCTTGGTCTGCTGCCCTAGATACAGAGTTTATACCCACAC 664
777 ACCTGTCTACACTGACATTCATTAAGTG-AGTGCTTGTGAAAGAAACAAATAAAC 835
665 ACCTGTCTACAGTGTCAATTAAGTGACGCTGCTTGTGAAAGAAACAAATAAAC 724
836 CC 837
725 CC 726

```

```

RESULT 3
BU603243
LOCUS
DEFINITION AGENCOURT 10059831 NIH MGC 142 Homo sapiens cdna clone
IMAGE: 6498263 5', mRNA sequence.
ACCESSION BU603243
VERSION BU603243.1 GI:23255002
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 835)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: NCI
CDNA Library Preparation: Michael Brownstein Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM2680 row: j column: 24
High quality sequence stop: 558.
Location/Qualifiers
1. 835
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6498263"
/tissue_type="mixed (pool of 40 RNAs)"
/lab_hosts="DH10B (TI-phage-resistant)"
/clone_lib="NIH_MGC_142"
/note="Vector: pDNR-LIB; Site 1: SfiI (ggcattatggcc);
Site 2: SfiI (ggcgctcgccg); Double-stranded cDNA was
prepared from a pool of 40 cell line polyA+ RNAs (bladder
- 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -
4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,
kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%,
ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary
gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were
used in cloning as follows:

```

FEATURES

```

source
1. 835
Query Match 61.8%; Score 685; DB 13; Length 835;
Best Local Similarity 96.6%; Pred. No. 1.6e-128;
Matches 705; Conservative 0; Mismatches 24; Indels 1; Gaps 1;
QY 106 AGCAGGAAATAGAACTTAAGAGAAATACACACTTCTGAGAAACTGAAACGACAGGGGAA 165
Db 8 AGCAGGAAATAGAACTTAAGAGAAATACACACTTCTGAGAAACTGAAACGACAGGGGAA 67
QY 166 AGGAGGTCTCACTGAGCACCGTCCAGCATCCGACACACACAGCGGCCCTTCCTCCACG 225
Db 68 AGGAGGTCTCACTGAGCACCGTCCAGCATCCGACACACAGCGGCCCTTCCTCCACG 127
QY 226 CAGAAACACACACTTCTCAAAACCTTCACTCAACACTTCTCCCAAGCAGGAAGATGC 285
Db 128 CAGAAACACACACTTCTCAAAACCTTCACTCAACACTTCTCCCAAGCAGGAAGATGC 187
QY 286 ACAAGGAGGACATGAGTGGCTGTGGGGGACCCCGAGCACCCTTCCTCCAGGT 345
Db 188 ACAAGGAGGACATGAGTGGCTGTGGGGGACCCCGAGCACCCTTCCTCCAGGT 247
QY 346 CCACCGTGATCAACATCCACAGCAGACCTCCGTGCCCGACCATGTGCTGTGCTCCCTGT 405
Db 248 CCACCGTGATCAACATCCACAGCAGACCTCCGTGCCCGACCATGTGCTGTGCTCCCTGT 307
QY 406 TCAACACCTCTTCTTGAACCTGCTGTGGGCTTCAATAGCATTCGCTACTCCGTGA 465
Db 308 TCAACACCTCTTCTTGAACCTGCTGTGGGCTTCAATAGCATTCGCTACTCCGTGA 367
QY 466 AGTCTAGGGACAGGAAGATGTTGGCGACGTGACCGGGGCCCGAGGCTATGCCTCCACCG 525
Db 368 AGTCTAGGGACAGGAAGATGTTGGCGACGTGACCGGGGCCCGAGGCTATGCCTCCACCG 427
QY 526 CCAAGTCCCTGAACATCTGGGCCCTGATCTGGGCATCTCTCATGACCATGGAATTCATCC 585
Db 428 CCAAGTCCCTGAACATCTGGGCCCTGATCTGGGCATCTCTCATGACCATGGAATTCATCC 487
QY 586 TGTACTGGTATTTCGGCTCTGTGACAGTCTACCATATTATGTTACAGATAAATACA-GGAA 644
Db 488 TGTACTGGTATTTCGGCTCTGTGACAGTCTACCATATTATGTTACAGATAAATACA-GGAA 547
QY 645 AAACGGGGTTACTAGTAGCCGCCCATAGCCTGCAACCTTTGCACTCCACTGTGCAATGCT 704
Db 548 AAACGGGGTTACTAGTAGCCGCCCATAGCCTGCAACCTTTGCACTCCACTGTGCAATGCT 607
QY 705 GGCCTGCACGCTGGGGCTGTGGCCCTGCCCTTGGTCTGCTGCCCTAGATACAGAGT 764
Db 608 GGCCTGCACGCTGGGGCTGTGGCCCTGCCCTTGGTCTGCTGCCCTAGATACAGAGT 667
QY 765 TTATACCCACACACCTGTCTACACTGACATTCATTAAGTGAAGTGAAGTGAAGTGAAGTGA 824
Db 668 TTATACCCACACACCTGTCTACAGTGTCAATTCATAAAGTGCAGCGCTTTGTGANNAA 727
QY 825 ACAATAAAAA 834
Db 728 AANNANNNNA 737

```

```

RESULT 4
CD174239
LOCUS CD174239 728 bp mRNA linear EST 19-MAY-2003
DEFINITION AGENCOURT_13976705 NIH_MGC_172 Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CD174239
VERSION CD174239.1 GI:30857594
KEYWORDS EST.

```

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 728)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs@mail.nih.gov
 Tissue Procurement: Dr. Jamie Thompson, University of WI
 cDNA Library Preparation: Gina Zastrow-Hayes
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: NDRM31 row: a column: 20
 High quality sequence start: 23
 High quality sequence stop: 660.
 High quality sequence stop: 660.
 Location/Qualifiers
 1..728
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="embryonic stem cells, WA01, passage 38"
 /lab_host="DH10B Tona"
 /clone_lib="NIH_MGC_172"
 /note="Vector: pDONR201; Site 1: attP2; Site 2: attP1;
 LIBR PRIMING - oligo dt; METHOD - full-length enriched;
 Embryonic Stem Cells H1; LIBR PROVIDER - Bradfield"

Query Match 61.6%; Score 682.8; DB 14; Length 728;
 Best Local Similarity 98.2%; Pred. No. 4.6e-128;
 Matches 701; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

Qy 125 AGAGAAATACACACTTCTGAGAACTGAACGACGAGGGAAGAGTCTCACTGAGCAC 184
 Db 15 AGTAGAAACCACTTCTGAGAACTGAACGACGAGGGAAGAGTCTCACTGAGCAC 74
 Qy 195 GTCCAGACGCGGACACACAGCGGCGCTTCGCTCCAGCAGCAACACACTTCTCA 244
 Db 75 GTCCAGACGCGGACACACAGCGGCGCTTCGCTCCAGCAGCAACACACTTCTCA 134
 Qy 245 AACCTTCACTCAACACTTCTCCCAAGCCAGAGATGCACAGGAGGAACTGAGGT 304
 Db 135 AACCTTCACTCAACACTTCTCCCAAGCCAGAGATGCACAGGAGGAACTGAGGT 194
 Qy 305 GCGTGTGCTGGGGACACCCCGCAGCACCATTCTCCAGGTCCACCGTGATCAACATCCA 364
 Db 195 GCGTGTGCTGGGGACACCCCGCAGCACCATTCTCCAGGTCCACCGTGATCAACATCCA 254
 Qy 365 CAGCGAGACTCGGTGCGCCGACCATGTCGTGCTGCTGCTTCAACACCTCTCTTGA 424
 Db 255 CAGCGAGACTCGGTGCGCCGACCATGTCGTGCTGCTGCTTCAACACCTCTCTTGA 314
 Qy 425 CTGGTCTGTCTGGGCTTCTAGCATTCGCGTACTCCGTAAGTCTAGGACAGGAAGT 484
 Db 315 CTGGTCTGTCTGGGCTTCTAGCATTCGCGTACTCCGTAAGTCTAGGACAGGAAGT 374
 Qy 485 GGTGCGGACGTCACCGGGCCGAGGCTATGCTCCACCGCCAAAGTCGCGAATCTG 544
 Db 375 GGTGCGGACGTCACCGGGCCGAGGCTATGCTCCACCGCCAAAGTCGCGAATCTG 434
 Qy 545 GCGCCGTGATTCGGGATCTCTATGACCATTTGGATTATCATCTGTACTGTTACGGTTC 604
 Db 435 GCGCCGTGATTCGGGATCTCTATGACCATTTGGATTATCATCTGTACTGTTACGGTTC 494
 Qy 605 TGTGACAGCTACCATATTATGTACAGATAATACAGAAAACGGGTTACTAGTACCC 664
 Db 495 TGTGACAGCTACCATATTATGTACAGATAATACAGAAAACGGGTTACTAGTACCC 554

Qy 665 GCCATAGCTGCAACCTTTGCACTCCACTGTCATGCTGGCCCTGCACCTGGGGCTG 724
 Db 555 GCCATAGCTGCAACCTTTGCACTCCACTGTCATGCTGGCCCTGCACCTGGGGCTG 614
 Qy 725 TTGCCCTTGGCCCTTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 784
 Db 615 TTGCCCTTGGCCCTTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 674
 Qy 785 ACAGTGACATTCATTAAGTG-ACGTGCTTGTGAAAAAACAATAAAACCC 837
 Db 675 ACAGTGATTCATTAAGTGACGCTGCTGTGAAAAAACAATAAAACCC 728

RESULT 5
 CF780624
 LOCUS DB 718 bp mRNA linear EST 20-OCT-2003
 DEFINITION AGENCOURT_15939275 NIH_MGC_217 Homo sapiens cDNA clone
 IMAGE:30524434 5', mRNA sequence.
 ACCESSION CF780624
 VERSION CF780624.1 GI:37740144
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 718)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgabbs@mail.nih.gov
 Tissue Procurement: James Martin, University of Iowa
 cDNA Library Preparation: M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: NDAM603 row: p column: 11
 High quality sequence stop: 626.
 High quality sequence stop: 626.
 Location/Qualifiers
 1..718
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30524434"
 /tissue_type="Pooled Chondrosarcoma Tumor cells"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NIH_MGC_217"
 /note="Vector: pYX-Asc; Site 1: EcoRI; Site 2: NotI;
 Library is oligo-dr primed and directionally cloned
 Denatured RNA was size fractionated on a 1% agarose gel.
 First strand cDNA synthesis was primed with oligo-dr
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated with
 EcoR I adaptor, digested with Not I and then cloned
 directionally into pYX-Asc vector. Average insert size
 0.5-1kb. Adaptors 5'(AATCGCAGGAGG)3' and 5'd
 (CTCTGTCGG)3'. 3' Linker sequence - GCGCGCTGAGAGCC T18.
 Sequencing primers 3'end: T3 promoter primer 5'd
 (ATTACCTCTCAAGGGA)3'. 5' End: T7 promoter primer 5'd
 (TAATACGACTACTAGGG)3'. Average insert size 0.5-1kb.
 Library was constructed in the laboratory of M. Bento
 Soares. Note: this is a NIH_MGC Library."

FEATURES
 source
 Query Match 61.5%; Score 682.2; DB 14; Length 718;
 Best Local Similarity 99.3%; Pred. No. 6e-128;
 Matches 706; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

ORIGIN

117 GAACTTAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAAAGAGGTCTCA 176
 Db 7 GAACTTAGAGAAATACACACTTCTGAGAACTGAAACGACAGGGGAAAGAGGTCTCA 66
 QY 177 CTGAGCACCGTCCAGCATCGGACACACAGCGGCCCTTCGCTCCACGACAGAAACACAC 236
 Db 67 CTGAGCACCGTCCAGCATCGGACACACAGCGGCCCTTCGCTCCACGACAGAAACAC 126
 QY 237 ACTTCTCAAACTTCACTCAACACTTCTTCCCAAAGCCGAGAGATGCAAGAGGAA 296
 Db 127 ACTTCTCAAACTTCACTCAACACTTCTTCCCAAAGCCGAGAGATGCAAGAGGAA 186
 QY 297 CATGAGGTGCTGCTGCTGGGGGACCCCGGACGACCATCTTCCAAAGTCCACCGTGATC 356
 Db 187 CATGAGGTGCTGCTGCTGGGGGACCCCGGACGACCATCTTCCAAAGTCCACCGTGATC 246
 QY 357 AACATCCACGAGAGACTCGTCCGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 416
 Db 247 AACATCCACGAGAGACTCGTCCGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 306
 QY 417 TTCTTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 476
 Db 307 TTCTTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 366
 QY 477 AGAAAGATGCTGCGGACGTCACCGGGGCCGAGGCTATGCTCCACCGCCCAAGTGCCTG 536
 Db 367 AGAAAGATGCTGCGGACGTCACCGGGGCCGAGGCTATGCTCCACCGCCCAAGTGCCTG 426
 QY 537 AACATCTGGGCGCTGATCTGGGATCTCATGACCATGATGATGATGATGATGATGATGATGAT 596
 Db 427 AACATCTGGGCGCTGATCTGGGATCTCATGACCATGATGATGATGATGATGATGATGATGAT 486
 QY 597 TTCGGCTCTGACAGTCTACCATATGTTTACAGATAATACAGGAAACGCGGTATAC 656
 Db 487 TTCGGCTCTGACAGTCTACCATATGTTTACAGATAATACAGGAAACGCGGTATAC 546
 QY 657 TAGTAGCGGCGCTGACAGTCTGCACTTTCGACTCCACTGTGCAATGCTGCGCTGCAACGC 716
 Db 547 TAGTAGCGCTCCATAGCTGCAACTTTCGACTCCACTGTGCAATGCTGCGCTGCAACGC 606
 QY 717 TGGGCGTGTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 776
 Db 607 TGGGCGTGTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 666
 QY 777 A-CCTGTCTACACTGACATTCATAAAGT-ACGTGCTTGTGAAAAA 825
 Db 667 ACCGTGCTACAGTGTCAATCAATAAGTGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 717

RESULT 6
 CD175008
 LOCUS
 DEFINITION
 AGENCOURT_13971378 NIH_MGC_172 Homo sapiens cDNA 5', mRNA sequence.
 CD175008.1 GI:30859148
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/
 1 (bases 1 to 717)
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Jamie Thompson, University of WI
 cDNA Library Preparation: Gina Zastrow-Hayes
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov

Plate: NDKM29 row: c column: 01
 High quality sequence stop: 717.
 Location/Qualifiers
 1. 717
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="embryonic stem cells, WA01, passage 38"
 /lab_host="DH10B Tona"
 /clone_lib="NIH_MGC_172"
 /note="Vector: pDONR201; Site 1: attP2; Site 2: attP1;
 LIBR PRIMING - oligo dT; METHOD - full-length enriched;
 Embryonic Stem Cells H1; LIBR PROVIDER - Bradfield"

ORIGIN
 Query Match 61.2%; Score 678.2; DB 14; Length 717;
 Best Local Similarity 98.0%; Pred. No. 3.9e-127; Indels 1; Gaps 1;
 Matches 697; Conservative 0; Mismatches 13;
 QY 128 GAAATACACACTTCTGAGAACTGAAACGACAGGGGAAAGAGGTCTCACTGAGCACCGT 187
 Db 6 GAAGGAGTAGCAACTGAGAACTGAAACGACAGGGGAAAGAGGTCTCACTGAGCACCGT 65
 QY 188 CCAGCATCCGACACACAGCGGCCCTTCGCTCCACGAGAAACACCACTTCTCAAC 247
 Db 66 CCAGCATCCGACACACAGCGGCCCTTCGCTCCACGAGAAACACCACTTCTCAAC 125
 QY 248 CTTCACTCAACACTTCTTCCCCAAAGCCAGAGATGCAAGAGGAAACATGAGGTGGC 307
 Db 126 CTTCACTCAACACTTCTTCCCCAAAGCCAGAGATGCAAGAGGAAACATGAGGTGGC 185
 QY 308 TGTGCTGGGGGACACCCCGGACCATCTTCCAAAGTCCACCGTATCAACATCCACAG 367
 Db 186 TGTGCTGGGGGACACCCCGGACCATCTTCCAAAGTCCACCGTATCAACATCCACAG 245
 QY 368 CGAGACCTCGGTGCGCGGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 427
 Db 246 CGAGACCTCGGTGCGCGGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 305
 QY 428 GTGCTGCTGGGCTTCATAGCATTCGCTTCTGAGTCTAGGAGACGAGATGGT 487
 Db 306 GTGCTGCTGGGCTTCATAGCATTCGCTTCTGAGTCTAGGAGACGAGATGGT 365
 QY 488 TGGGAGTGAACCGGGGCCAGGCTATGCTCCACCGGCAAGTGCCTGACATCTGGG 547
 Db 366 TGGGAGTGAACCGGGGCCAGGCTATGCTCCACCGGCAAGTGCCTGACATCTGGG 425
 QY 548 CCTGATCTGGGCTATCCTCATGACCATGGATTCATCTGTTACTGCTATTCGCTCTGT 607
 Db 426 CCTGATCTGGGCTATCCTCATGACCATGGATTCATCTGTTACTGCTATTCGCTCTGT 485
 QY 608 GACAGTCTACATATATATGTTACAGATAATACAGGAAACCGGGGTACTAGTACCGCC 667
 Db 486 GACAGTCTACATATATATGTTACAGATAATACAGGAAACCGGGGTACTAGTACCGCC 545
 QY 668 CATAGCTGCAACCTTTGCACTGCCATGCTGCAATGCTGCGCTGCGCTGCGCTGCTGCTGCTGCT 727
 Db 546 CATAGCTGCAACCTTTGCACTGCCATGCTGCAATGCTGCGCTGCGCTGCGCTGCTGCTGCTGCT 605
 QY 728 CCCCTGCCCTTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 787
 Db 606 CCCCTGCCCTTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 665
 QY 788 CTGACATTCATAAAGTG-ACGTGCTTGTGAAAAAACAATAAATACCC 837
 Db 666 GTGTCATTCATAAAGTGACGTGCTTGTGAAAAAACAATAAATACCC 716

RESULT 7
 CD175837
 LOCUS
 DEFINITION
 AGENCOURT_13976863 NIH_MGC_172 Homo sapiens cDNA 5', mRNA sequence.
 CD175837
 ACCESSION

VERSION	CD175837.1	GI:30860834
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 726)	
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.	
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)	
JOURNAL	Unpublished (1999)	
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs@mail.nih.gov Tissue Procurement: Dr. Jamie Thompson, University of WI cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: NDKM51 row: m column: 03 High quality sequence start: 16 High quality sequence stop: 718.	
FEATURES	Location/Qualifiers	
source	1..726	
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/tissue_type="embryonic stem cells, WA01, passage 38"	
	/lab_host="DH10B Tona"	
	/clone_lib="NIH MGC 172"	
	/note="Vector: pDONR201; Site 1: attP2; Site 2: attP1; LIBR PRIMING - oligo dt; METHOD - full-length enriched; Embryonic Stem Cells H1; LIBR PROVIDER - Bradfield"	
ORIGIN		
Query Match	61.1%; Score 678; DB 14; Length 726;	
Best Local Similarity	97.5%; Pred. No. 4.3e-127;	
Matches	698; Conservative 0; Mismatches 17; Indels 1; Gaps 1;	
QY	120 ACTTAAAGAGAAATACACATCTCTGAGAACTGAAACGACAGGGGAAAGGAGGTCTCACTG 179	
DB	4 ACCGAAGAGTAGAAAACTCTGAGAACTGAAACGACAGGGGAAAGGAGGTCTCACTG 63	
QY	180 AGCACCGTCCAGCATCCCGACACACACAGCGGCCCTTCGCTCCACGACAGAAACCACT 239	
DB	64 AGCACCGTCCAGCATCCCGACACACACAGCGGCCCTTCGCTCCACGACAGAAACCACT 123	
QY	240 TCTCAAACTTCACTCAACATCTCTTCCCAAGCCAGAGATGCAAGGAGGAAACAT 299	
DB	124 TCTCAAACTTCACTCAACATCTCTTCCCAAGCCAGAGATGCAAGGAGGAAACAT 183	
QY	300 GAGGTGGTGTGTGGGGCACCACCCAGCACCATCTTCCAGAGTCCACCGTGATCAAC 359	
DB	184 GAGGTGGTGTGTGGGGCACCACCCAGCACCATCTTCCAGAGTCCACCGTGATCAAC 243	
QY	360 ATCCACAGCAGCATCTCCGTGCCGACCAATGTGCTGTGGTCCCTGTTCACACCCCTTTC 419	
DB	244 ATCCACAGCAGCATCTCCGTGCCGACCAATGTGCTGTGGTCCCTGTTCACACCCCTTTC 303	
QY	420 TTGAATGGTGTGTGGGTTCATAGCATTCGCTACTCCGTTGAAGTCTTAGGGACAGG 479	
DB	304 TTGAATGGTGTGTGGGTTCATAGCATTCGCTACTCCGTTGAAGTCTTAGGGACAGG 363	
QY	480 AAGATGGTTGGGACAGTGCAGCGGGGCCAGGCGCTATGCTCCACCGCAAGTGCCTGAAC 539	
DB	364 AAGATGGTTGGGACAGTGCAGCGGGGCCAGGCGCTATGCTCCACCGCAAGTGCCTGAAC 423	
QY	540 ATCTGGGCCCTGATTTCTGGGCATCTTCATGACATTTGATTCATCTCTGTTACTGATTC 599	
DB	424 ATCTGGGCCCTGATTTCTGGGCATCTTCATGACATTTGATTCATCTCTGTTACTGATTC 483	
QY	600 GGCTGTGTGACAGTCTACCATATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAG 659	
Db	484 GGCTGTGTGACAGTCTACCATATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAG 543	
QY	660 TAGCGCCCATAGCCTGCAACCTTTTGCACTCCACTGCTGCAATGCTGCGCCCTGCACTGG 719	
Db	544 TAGCGCCCATAGCCTGCAACCTTTTGCACTCCACTGCTGCAATGCTGCGCCCTGCACTGG 603	
QY	720 GGCTGTGGCCCTGCGCCCTTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 779	
Db	604 GGCTGTGGCCCTGCGCCCTTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 663	
QY	780 TGTCTACACTGACATTCATATAAAGTG-AGTGTGCTTGTGAAAAAACAATAAAAA 834	
Db	664 TGTCTACAGTGTCAATTCATATAAAGTGACAGTGTGCTTGTGAAAAAACAATAAAAA 719	
RESULT 8		
LOCUS	CD172759 718 bp mRNA linear EST 19-MAY-2003	
DEFINITION	AGENCOURT_13966328 NIH_MGC_172 Homo sapiens cDNA 5', mRNA sequence.	
ACCESSION	CD172759	
VERSION	CD172759.1	
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 718)	
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.	
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)	
JOURNAL	Unpublished (1999)	
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs@mail.nih.gov Tissue Procurement: Dr. Jamie Thompson, University of WI cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: NDKM28 row: d column: 04 High quality sequence start: 11 High quality sequence stop: 713.	
FEATURES	Location/Qualifiers	
source	1..718	
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/tissue_type="embryonic stem cells, WA01, passage 38"	
	/lab_host="DH10B Tona"	
	/clone_lib="NIH MGC 172"	
	/note="Vector: pDONR201; Site 1: attP2; Site 2: attP1; LIBR PRIMING - oligo dt; METHOD - full-length enriched; Embryonic Stem Cells H1; LIBR PROVIDER - Bradfield"	
ORIGIN		
Query Match	61.0%; Score 676.2; DB 14; Length 718;	
Best Local Similarity	98.6%; Pred. No. 9.9e-127;	
Matches	692; Conservative 0; Mismatches 9; Indels 1; Gaps 1;	
QY	125 AGAGAAATACACATCTCTGAGAACTGAAACGACAGGGGAAAGGAGGTCTCACTGAGCAC 184	
DB	15 AGTAGAAACACATCTCTGAGAACTGAAACGACAGGGGAAAGGAGGTCTCACTGAGCAC 74	
QY	185 COTCCAGCATCGGACACACAGCGGCCCTTCGCTCCAGCAGAAAAACCACTTCTCA 244	
DB	75 COTCCAGCATCGGACACACAGCGGCCCTTCGCTCCAGCAGAAAAACCACTTCTCA 134	
QY	245 AACCTTCACTCAACACTTCTCTCCCAAGCCAGAGATGCACAGGAGGAAACATGAGGT 304	
DB	135 AACCTTCACTCAACACTTCTCTCCCAAGCCAGAGATGCACAGGAGGAAACATGAGGT 194	
QY	305 GGCTGTGTGGGGGACCCCGCCAGCACCATCTCTTCCAGGTCCACCGTGTGATCAATCCA 364	

Db 195 TCGTGTGCGGGGACCCGCCAGCACCATCCCTCCAAAGTCCACCGTGATCAACATCCA 254
 QY 365 CAGCAGAGACTCGGTGCCGCCGACCATGTCGTCTCCCTGTTCAACACCCCTCTTCTTCAA 424
 Db 255 CAGCAGAGACTCGGTGCCGCCGACCATGTCGTCTCCCTGTTCAACACCCCTCTTCTTCAA 314
 QY 425 CTGGTCTGTCTGGGCTTCTATAGCATTCGCTACTCCGTGAGTCTAGGGACAGGAAGAT 484
 Db 315 CTGGTCTGTCTGGGCTTCTATAGCATTCGCTACTCCGTGAGTCTAGGGACAGGAAGAT 374
 QY 485 GGTGTGGCAGCTGACCGGGGCCAGGGGCTATGCTCCACCGCCCAAGTGCCTGAAACATCTG 544
 Db 375 GGTGTGGCAGCTGACCGGGGCCAGGGGCTATGCTCCACCGCCCAAGTGCCTGAAACATCTG 434
 QY 545 GGGCTGTATCTGGGATCTCTATGACCATTTGATGATTCATCTGTTACTGTTATTCGGCTC 604
 Db 435 GGGCTGTATCTGGGATCTCTATGACCATTTGATGATTCATCTGTTACTGTTATTCGGCTC 494
 QY 605 TGTGACAGTCTACCATATTTATGTTACAGATTAATACAGGAAAAACGGGGTTACTAGTAGCC 664
 Db 495 TGTGACAGTCTACCATATTTATGTTACAGATTAATACAGGAAAAACGGGGTTACTAGTAGCC 554
 QY 665 GGCATAGCTGCAACCTTTCGACTCCACTGTGCAATGCTGGCCCTGCACTGCGTGGGGTG 724
 Db 555 GGCATAGCTGCAACCTTTCGACTCCACTGTGCAATGCTGGCCCTGCACTGCGTGGGGTG 614
 QY 725 TTGCCCTGCCCCCTTGGTCTGCTCCCTAGTACAGAGTTATATACCCACACCTGTCT 784
 Db 615 TTGCCCTGCCCCCTTGGTCTGCTCCCTAGTACAGAGTTATATACCCACACCTGTCT 674
 QY 785 ACACTGACATTTCAATAAGTG-ACGTGCTTGTGAAAAA 825
 Db 675 ACAGTGTCTTCAATAAGTGACGTGCTTGTGAAAAA 716

RESULT 9
 CD176014
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 726)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-1@mail.nih.gov
 Tissue Procurement: Dr. Jamie Thompson, University of WI
 CDNA Library Preparation: Gina Zastrow-Hayes
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: NDKM49 row: h column: 05
 High quality sequence start: 29
 High quality sequence stop: 696.
 Location/Qualifiers
 1. 726
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="embryonic stem cells, WA01, passage 38"
 /lab_host="DH10S Tera"
 /clone_lib="NIH_MGC_172"
 /note="Vector: pDONR201; Site 1: attP2; Site 2: attP1;
 LIBR PRIMING - oligo dt; METHOD - full-length enriched;
 Embryonic Stem Cells H1; LIBR PROVIDER - Bradfield"

ORIGIN
 Query Match 60.7%; Score 672.8; DB 14; Length 726;
 Best Local Similarity 98.9%; Pred. No. 4.8e-126;
 Matches 698; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
 QY 143 GAGAACTGAACGACAGAGGGGAAAGGAGGTCTCACTGAGCAGCCGTCCTCCAGCATCCGACA 202
 Db 24 GAGAACTCTAAACGACAGAGGGGAAAGGAGGTCTCACTGAGCAGCCGTCCTCCAGCATCCGACA 83
 QY 203 CCACAGCGGCCCTTCGCTCCACGAGAAAAACCACTTCTCAAACTTCTCAACCTTCAACACTT 262
 Db 84 CCACAGCGGCCCTTCGCTCCACGAGAAAAACCACTTCTCAAACTTCTCAACACTTCAACACTT 143
 QY 263 CTTTCCCCAAGCAGAGAGATGACAGAGGAGGACATGAGGTGGCTGTCTGGGGCACC 322
 Db 144 CTTTCCCCAAGCAGAGAGATGACAGAGGAGGACATGAGGTGGCTGTCTGGGGCACC 203
 QY 323 CCCAGCACCATCTTCCAAAGGTCCACCGTGATCAACATCCACAGCAGACCTCCGTGCC 362
 Db 204 CCCAGCACCATCTTCCAAAGGTCCACCGTGATCAACATCCACAGCAGACCTCCGTGCC 263
 QY 383 GACCATGTGCTGTGTCCTTCAACACCTCTTCTTGAACCTGGTGTCTGGGCTT 442
 Db 264 GACCATGTGCTGTGTCCTTCAACACCTCTTCTTGAACCTGGTGTCTGGGCTT 323
 QY 443 CATAGCATTCGCTTACTCCGTTAGTCTAGGACAGGAAGATGGTTGGCGAAGTGAACCGG 502
 Db 324 CATAGCATTCGCTTACTCCGTTAGTCTAGGACAGGAAGATGGTTGGCGAAGTGAACCGG 383
 QY 503 GGCCCGAGGCTATGCTCCACCGCAAGTGGCTGAACATCTGGGCGCTGATTTCTGGGCAT 562
 Db 384 GGCCCGAGGCTATGCTCCACCGCAAGTGGCTGAACATCTGGGCGCTGATTTCTGGGCAT 443
 QY 563 CCTCATGACCATTTGATTCCTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 622
 Db 444 CCTCATGACCATTTGATTCCTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 503
 QY 623 TATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGCCCATAGCTGCAACCT 682
 Db 504 TATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGCCCATAGCTGCAACCT 563
 QY 683 TTGACTTCCACTGTGCAATGCTGGCCCTGCACTGGGGCTGTGGCCCTGCCCTTGG 742
 Db 564 TTGACTTCCACTGTGCAATGCTGGCCCTGCACTGGGGCTGTGGCCCTGCCCTTGG 623
 QY 743 TCCTGCCCTTAGATACAGCAGCTTTATACCCACACACCTGTCTACACTGACATTTCAATAAA 802
 Db 524 TCCTGCCCTTAGATACAGCAGCTTTATACCCACACCTGTCTACACTGACATTTCAATAAA 683
 QY 803 GTG-ACGTGCTTGTGAAAAAATAAATAAACC 837
 Db 684 GTGACCGTCTTGTGAAAAAATAAATAAACC 719
 RESULT 10
 CD175245
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 721)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-1@mail.nih.gov
 Tissue Procurement: Dr. Jamie Thompson, University of WI

cdna Library Preparation: Gina Zastrow-Hayes
cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDKM41 row: C column: 22
High quality sequence start: 4
High quality sequence stop: 721.
Location/Qualifiers

FEATURES

1..721
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, WA01, passage 38"
/lab_host="DHIOB T0A"
/clone_lib="NIH_MGC_172"
/notes="Vector: pDONR201; Site 1: attP2; Site 2: attP1;
LIBR PRIMING - oligo dT; METHOD - full-length enriched;
Embryonic Stem Cells H1; LIBR PROVIDER - Bradfield"

ORIGIN

Query Match 60.6%; Score 671.8; DB 14; Length 721;
Best Local Similarity 98.8%; Pred. No. 7.7e-126;
Matches 687; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
QY 141 CTGAGAACTGAAACGACAGCGGGAAGAGGTCTCACTGAGCAGCGTCCAGCATCCGGA 200
DB 20 CTGAGAACTGAAACGACAGCGGGAAGAGGTCTCACTGAGCAGCGTCCAGCATCCGGA 79
QY 201 CACACAGCGGCGCTTGGTCCACGCGAGAAACACACACTTCTCAAACTTCACCTCAACAC 260
DB 80 CACACAGCGGCGCTTGGTCCACGCGAGAAACACACACTTCTCAAACTTCACCTCAACAC 139
QY 261 TTCCTTCCCAAGCCAGAGATGCAAGAGGAGCAATGAGTGGCTGTCTGGGGCA 320
DB 140 TTCCTTCCCAAGCCAGAGATGCAAGAGGAGCAATGAGTGGCTGTCTGGGGCA 199
QY 321 CCCCCAGCACCATCTCTCAAGGTCCACCGTGATCAACATCCACAGCAGACCTCCCGT 380
DB 200 CCCCCAGCACCATCTCTCAAGGTCCACCGTGATCAACATCCACAGCAGACCTCCCGT 259
QY 381 CCGGACCATGTGCTGCTGCTTCAACACCTCTTGTGACTGCTGTCTGGG 440
DB 260 CCGGACCATGTGCTGCTGCTTCAACACCTCTTGTGACTGCTGTCTGGG 319
QY 441 TTCATAGCATTCGCTTACTCCGTGAAGTCTAGGACAGGAAGTGGTGGGAGTGACC 500
DB 320 TTCATAGCATTCGCTTACTCCGTGAAGTCTAGGACAGGAAGTGGTGGGAGTGACC 379
QY 501 GGGGCCCAGGCTATGCTCCACGCGCAAGTGCCTGAACATCTGGGCGCTGATCTGGG 560
DB 380 GGGGCCCAGGCTATGCTCCACGCGCAAGTGCCTGAACATCTGGGCGCTGATCTGGG 439
QY 561 ATCTCTCATGACCATTCGATTCATCTCTTACTGTATTCGCTCTGTGACAGTCTACCAT 620
DB 440 ATCTCTCATGACCATTCGATTCATCTCTTACTGTATTCGCTCTGTGACAGTCTACCAT 499
QY 621 ATTATGTTACAGATAATACAGGAAAAACGGGTTACTAGTAGCGCCCATAGCTTCAAC 680
DB 500 ATTATGTTACAGATAATACAGGAAAAACGGGTTACTAGTAGCGCCCATAGCTTCAAC 559
QY 681 CTTTCGACTCCAGTGTGCAATGCTGGCCCTGACGCTGGGCTGTTCCTGCGCCCTT 740
DB 560 CTTTCGACTCCAGTGTGCAATGCTGGCCCTGACGCTGGGCTGTTCCTGCGCCCTT 619
QY 741 GGTCTCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACACTGACATTCATA 800
DB 620 GGTCTCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATTCATA 679
QY 801 AAGTG-ACGTGCTTGTGAAAAAATAAATAA 834
DB 680 AAGTGACGTGCTTGTGAAAAAATAAATAA 714

RESULT 11

AV759074
LOCUS AV759074 MDS Homo sapiens cdna clone MUSEPH09 5', mRNA sequence.
DEFINITION AV759074 MDS Homo sapiens cdna clone MUSEPH09 5', mRNA sequence.
ACCESSION AV759074
VERSION AV759074.1 GI:10916922
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 721)
AUTHORS Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H.,
Gu, Y., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z.,
Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M.,
Lu, G., Yang, Y., Gao, G., Zhang, Q., Chen, S., Han, Z. and Chen, Z.
TITLE Homo sapiens cdna MDS clones
JOURNAL Unpublished (2000)
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES

source
1..727
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MUSEPH09"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="MDS"
/notes="Vector: pTriplex2; Site 1: sfiIA; Site 2: sfiIB"

ORIGIN

Query Match 60.3%; Score 668.8; DB 9; Length 727;
Best Local Similarity 98.4%; Pred. No. 3.1e-125;
Matches 695; Conservative 0; Mismatches 9; Indels 2; Gaps 2;
QY 129 AAATACACACTTCTGAGAA-ACTGAAACGACAGGGGAAAGGAGGTCTCACTGAGCACCGT 187
DB 1 AAATACACACTTCTGAGAGACTGAAACGACAGGGGAAAGGAGGTCTCACTGAGCACCGT 60
QY 188 CCGAGATCCGACACACACAGCGGCGCTTCGCTCCACGAGAAACCACTTCTCAAC 247
DB 61 CCGAGATCCGACACACACAGCGGCGCTTCGCTCCACGAGAAACCACTTCTCAAC 120
QY 248 CTTCACTCAACACTTCTCTCCCAAGCCAGAGATGCAAGGAGGAAACATGAGGTGGC 307
DB 121 CTTCACTCAACACTTCTCTCCCAAGCCAGAGATGCAAGGAGGAAACATGAGGTGGC 180
QY 308 TGTGCTGGGGGACACCCCGGACCATCTCTTCCAAAGTCCACCGTGTCAACATCCACAG 367
DB 181 TGTGCTGGGGGACACCCCGGACCATCTCTTCCAAAGTCCACCGTGTCAACATCCACAG 240
QY 368 CGAGACCTCCGTCGCGCGACCATGCTGCTGCTGCTTCAACACCCCTCTTCTTCAACTG 427
DB 241 CGAGACCTCCGTCGCGCGACCATGCTGCTGCTGCTTCAACACCCCTCTTCTTCAACTG 300
QY 428 GTGCTGTCTGGGCTTCATAGCATTCGCTACTCCGTAAGTCTAGGGACAGGAAGATGGT 487
DB 301 GTGCTGTCTGGGCTTCATAGCATTCGCTACTCCGTAAGTCTAGGGACAGGAAGATGGT 360
QY 488 TGGCGACGTGACCGGGGCGGCGCTATGCTCCACCGCAAGTGCCTGAACATCTGGGC 547
DB 361 TGGCGACGTGACCGGGGCGGCGCTATGCTCCACCGCAAGTGCCTGAACATCTGGGC 420

```
QY 548 CTTGATTCTGGGATCCTCATGACCAATGGATTATCTCTGTTACTGCTGTTATTCGGCTCTGT 607
Db 421 CCTGATTCTGGGATCCTCATGACCAATGGATTATCTCTGTTACTGCTGTTATTCGGCTCTGT 480
QY 608 GACAGCTCTACCATATTATGTTACAGATAATACAGGAAAAACGGGGTACTAGTAGCGGCC 667
Db 481 GACAGCTCTACCATATTATGTTACAGATAATACAGGAAAAACGGGGTACTAGTAGCGGCC 540
QY 668 CATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTGCAATGCTGCAATGCTGCAATGCTGCAAT 727
Db 541 CATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTGCAATGCTGCAATGCTGCAATGCTGCAAT 600
QY 728 CCCTGCCCCCTTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 787
Db 601 CCCTGCCCCCTTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 788 CTGACATTCAATAAGTG-ACGTGCTTGTGAAAAAACAATAA 832
Db 661 GTGTCATTCAATAAGTGACGTGCTTGTGAAAAAACAATAA 706

RESULT 12
CD173064
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CD173064 704 bp mRNA linear EST 19-MAY-2003
AGENCOURT_13974182 NIH_MGC_172 Homo sapiens cDNA 5', mRNA sequence.
CD173064.1 GI:30855257
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Tissue Procurement: Dr. Jamie Thompson, University of WI
cDNA Library Preparation: Gina Zastrow-Hayes
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDKM46 row: h column: 04
High quality sequence start: 4
High quality sequence stop: 704.
High quality sequence stop: 704.
Location/Qualifiers
1. 704
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, WA01, passage 38"
/lab_host="DH10B Tona"
/clone_lib="NIH-MGC 172"
/notes="vector: pDONR201; Site 1: attP2; Site 2: attP1;
LIBR PRIMING - oligo dt; METHOD - full-length enriched;
Embryonic Stem Cells H1; LIBR PROVIDER - Bradfield"

ORIGIN
Query Match 60.2%; Score 668; DB 14; Length 704;
Best Local Similarity 98.4%; Pred. No. 4.5e-125;
Matches 68; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

QY 143 GAGAACTGAACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 202
Db 9 GAGTAGAAGAACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 68
QY 203 CCACAGGGCCCTTGGCTCCAGCGGAGAGAAACACACTTCTCAACCTTCACTCAACACTT 262
Db 69 CCACAGGGCCCTTGGCTCCAGCGGAGAGAAACACACTTCTCAACCTTCACTCAACACTT 128
```

```
QY 263 CTTTCCCAAGACCAGAGATGCAAGAGAAATGAGAGTGGCTGTGTCTGGGGGCAAC 322
Db 129 CTTTCCCAAGACCAGAGATGCAAGAGAAATGAGAGTGGCTGTGTCTGGGGGCAAC 188
QY 323 CCCAGACACCATCTTCCAAAGGTCCACCGTGATCAACATCCACAGCGAGACCTCCGTCGCC 382
Db 189 CCCAGACACCATCTTCCAAAGGTCCACCGTGATCAACATCCACAGCGAGACCTCCGTCGCC 248
QY 383 CGACCATGTGCTCTGGTCCCTGTTCAACACCCCTCTTCTTGAATGCTGTCTGGGCTT 442
Db 249 CGACCATGTGCTCTGGTCCCTGTTCAACACCCCTCTTCTTGAATGCTGTCTGGGCTT 308
QY 443 CATAGCATTTGSCCTTACTCCGTGAGTCTAGGACAGGAGAGATGGTTGGGAGACGTGACCGG 502
Db 309 CATAGCATTTGSCCTTACTCCGTGAGTCTAGGACAGGAGAGATGGTTGGGAGACGTGACCGG 368
QY 503 GGCCAGGCGCTATGCTCCACCGCAAGTGCTGGAACATCTGGGCGCTGATTTGGGCGAT 562
Db 369 GGCCAGGCGCTATGCTCCACCGCAAGTGCTGGAACATCTGGGCGCTGATTTGGGCGAT 428
QY 563 CCTCATGACCATTTGGATTCTATCCTGTTACTGTTATTCGGCTCTGTGACAGTCTACCATAT 622
Db 429 CCTCATGACCATTTGGATTCTATCCTGTTACTGTTATTCGGCTCTGTGACAGTCTACCATAT 488
QY 623 TATGTTACAGATAATACAGGAAAAACGGGGTACTAGTAGCGGCCCATAGCCTGCAACCT 682
Db 489 TATGTTACAGATAATACAGGAAAAACGGGGTACTAGTAGCGGCCCATAGCCTGCAACCT 548
QY 683 TTGACATCCACTGTGCAATGCTGCGCTGCACGCTGGGGCTGTGCGCTGCCCTTGG 742
Db 549 TTGACATCCACTGTGCAATGCTGCGCTGCACGCTGGGGCTGTGCGCTGCCCTTGG 608
QY 743 TCCTGCCCCCTAGATACAGCAGTTTATACCCACACACACCTGTCTACAGTGCATTCATATAA 802
Db 609 TCCTGCCCCCTAGATACAGCAGTTTATACCCACACACACCTGTCTACAGTGCATTCATATAA 668
QY 803 GTG-ACGTGCTTGTGAAAAAACAATAAATACCC 837
Db 669 GTGACGTGCTTGTGAAAAAACAATAAATACCC 704

RESULT 13
CD175397
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CD175397 715 bp mRNA linear EST 19-MAY-2003
AGENCOURT_13971659 NIH_MGC_172 Homo sapiens cDNA 5', mRNA sequence.
CD175397.1 GI:30859931
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Tissue Procurement: Dr. Jamie Thompson, University of WI
cDNA Library Preparation: Gina Zastrow-Hayes
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDKM41 row: n column: 17
High quality sequence start: 2
High quality sequence stop: 715.
High quality sequence stop: 715.
Location/Qualifiers
1. 715
/organism="Homo sapiens"
/db_xref="taxon:9606"

FEATURES
source
```



```

Db      618 TCCCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCTATTCAATAGTG 677
QY      806 -ACGTGCTTTGTGAAAAAACAATAAAAACC 837
Db      678 CACGTGCTTGTGAAAAAACAATAAAAACC 710

RESULT 15
CD175286
LOCUS   CD175286              712 bp      mRNA      linear      EST 19-MAY-2003
DEFINITION AGENCOURT_13978429 NIH_MGC_172 Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CD175286
VERSION   CD175286.1 GI:30859713
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1. (bases 1 to 712)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Dr. Jamie Thompson, University of WI
cDNA Library Preparation: Gina Zastrow-Hayes
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: Agencourt Bioscience Corporation
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDKM49 row: m column: 06
High quality sequence start: 31
High quality sequence stop: 712.
Location/Qualifiers
1. 712
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, WA01, passage 38"
/lab_host="DH10B Tona"
/clone_lib="NIH MGC 172"
/note="Vector: pDONR201; Site 1: attP2; Site 2: attP1;
LIBR PRIMING - oligo dT; METHOD - full-length enriched;
Embryonic Stem Cells H1; LIBR PROVIDER - Bradfield"

FEATURES
source

ORIGIN
Query Match 60.1%; Score 666.4; DB 14; Length 712;
Best Local Similarity 98.3%; Pred.No. 9.6e-125;
Matches 684; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

QY      143 GAGAACTGAAACGACGAGGGAAGAGGTCTCACTGAGCAGCAGTCCGACATCCGGACA 202
Db      17  GAGTAGAAAAACGACGAGGGAAGAGGTCTCACTGAGCAGCAGTCCGACATCCGGACA 76
QY      203 CCACAGCGCCCTTCGCTCCAGCAGAAAAACCACTTCTCAAACTTCACTCAACACTT 262
Db      77  CCACAGCGCCCTTCGCTCCAGCAGAAAAACCACTTCTCAAACTTCACTCAACACTT 136
QY      263 CTTTCCCAAGCCAGAAATGACAGAGGAGAACATGAGTGGCTGTGTGGGGGACC 322
Db      137 CTTTCCCAAGCCAGAAATGACAGAGGAGAACATGAGTGGCTGTGTGGGGGACC 196
QY      323 CCCAGCACCATCTTCCAGGTCCACCGTGATCAACATCCAGCAGAGACCTCCGTGCC 382
Db      197 CCCAGCACCATCTTCCAGGTCCACCGTGATCAACATCCAGCAGAGACCTCCGTGCC 256
QY      383 CGACCATGTGCTGCTGCTGCTTCAACACCTTCTTGAACCTGTGTGTGGGCTT 442
Db      257  CGACCATGTGCTGCTGCTGCTTCAACACCTTCTTGAACCTGTGTGTGGGCTT 316
QY      443 CATAGCATTCGCTACTCCGTGAAGTCTAGGGACAGGAAGATGGTTGGCGACGTGACCGG 502

```

```

Db      317 CATAGCATTCGCTACTCCGTGAAGTCTAGGACAGGAAGATGGTTGGCAGCTGACCGG 376
QY      503 GGGCCAGGCGCTATGCTCCACCGCCAAAGTGCCTGAACATCTGGGCGCTGATTTCTGGGCAT 562
Db      377 GGGCCAGGCGCTATGCTCCACCGCCAAAGTGCCTGAACATCTGGGCGCTGATTTCTGGGCAT 436
QY      563 CCTCATGACCATTTGGATTCTATCTCTGTTACTGGTATTCTGGCTCTGTGACAGTCTACCATAT 622
Db      437 CTTTATGACCATTTGGATTCTATCTCTGTTACTGGTATTCTGGCTCTGTGACAGTCTACCATAT 496
QY      623 TATGTTACAGATAATAACAGAAAAACGGGGTTACTAGTAGCCGCCCATAGCCTGCAACCT 682
Db      497 TATGTTACAGATAATAACAGAAAAACGGGGTTACTAGTAGCCGCCCATAGCCTGCAACCT 556
QY      683 TTGCACTCCACTGTGCAATGCTGCCCTGCAAGTGGGGTGTGGCCCTGCCCCCTTGG 742
Db      557 TTGCACTCCACTGTGCAATGCTGCCCTGCAAGTGGGGTGTGGCCCTGCCCCCTTGG 616
QY      743 TCTTGGCCCTTAGATACAGCAGTTTATACCCACACACACCTGTCTACACTGACATTTCAATAAA 802
Db      617 TCTTGGCCCTTAGATACAGCAGTTTATACCCACACACACCTGTCTACAGTGTCTCAATAAA 676
QY      803 GTG-ACGTGCTTGTGAAAAAACAATAAAAACC 837
Db      677 GTGACAGTGTGTTGAAAAAACAATAAAAACC 712

```

Search completed: March 11, 2004, 14:23:57

Job time : 1840.29 secs